(1)

Technical Report 799

Knowledge Elicitation of Recognition-Primed Decision Making

Gary A. Klein
Klein Associates, Inc.

Donald MacGregor
MacGregor-Bates, Inc.

ARI Field Unit at Fort Leavenworth, Kansas
Systems Research Laboratory

U. S. Army
Research Institute for the Behavioral and Social Sciences

July 1988

U. S. ARMY RESEARCH INSTITUTE

FOR THE BEHAVIORAL AND SOCIAL SCIENCES

A Field Operating Agency under the Jurisdiction of the

Deputy Chief of Staff for Personnel

EDGAR M. JOHNSON
Technical Director

WM. DARRYL HENDERSON
COL, IN
Commanding

Research accomplished under contract
for the Department of the Army

Klein Associates, Inc.

Technical review by

George W. Lawton
Robert Solick

NOTICES

~~DISTRIBUTION: Primary distribution of this report has been made by ARI. Please address correspondence concerning distribution of reports to: U.S. Army Research Institute for the Behavioral and Social Sciences, ATTN: PERI-POT, 5001 Eisenhower Ave., Alexandria, Virginia 22333-5600~~

FINAL DISPOSITION. This report may be destroyed when it is no longer needed. Please do not return it to the U.S. Army Research Institute for the Behavioral and Social Sciences.

NOTE. The findings in this report are not to be construed as an official Department of the Army position, unless so designated by other authorized documents.

REPORT DOCUMENTATION PAGE

1a. REPORT SECURITY CLASSIFICATION: Unclassified	1b. RESTRICTIVE MARKINGS
2a. SECURITY CLASSIFICATION AUTHORITY	3. DISTRIBUTION/AVAILABILITY OF REPORT: Approved for public release; distribution unlimited.
2b. DECLASSIFICATION/DOWNGRADING SCHEDULE	
4. PERFORMING ORGANIZATION REPORT NUMBER(S): KA-TR-863-87-06-F	5. MONITORING ORGANIZATION REPORT NUMBER(S): ARI Technical Report 799
6a. NAME OF PERFORMING ORGANIZATION: Klein Associates, Inc. / 6b. OFFICE SYMBOL (If applicable)	7a. NAME OF MONITORING ORGANIZATION: U.S. Army Research Institute Field Unit
6c. ADDRESS (City, State, and ZIP Code): P.O. Box 264, 800 Livermore Street, Yellow Springs, OH 45387	7b. ADDRESS (City, State, and ZIP Code): P.O. Box 3407, Fort Leavenworth, KS 66027-0347
8a. NAME OF FUNDING/SPONSORING ORGANIZATION: U.S. Army Research Institute / 8b. OFFICE SYMBOL (If applicable): PERI-BR	9. PROCUREMENT INSTRUMENT IDENTIFICATION NUMBER: MDA903-86-C-0170
8c. ADDRESS (City, State, and ZIP Code): 5001 Eisenhower Avenue, Alexandria, VA 22333-5600	10. SOURCE OF FUNDING NUMBERS: PROGRAM ELEMENT NO. 6.55.02.A; PROJECT NO. 2P665 502M770; TASK NO. 1.4.4; WORK UNIT ACCESSION NO. S.6

11. TITLE (Include Security Classification)

Knowledge Elicitation of Recognition-Primed Decision Making

12. PERSONAL AUTHOR(S)

Gary A. Klein and Donald MacGregor (MacGregor-Bates, Inc.)

13a. TYPE OF REPORT	13b. TIME COVERED	14. DATE OF REPORT (Year, Month, Day)	15. PAGE COUNT
Final	FROM 08/86 TO 09/87	1988, July	61

16. SUPPLEMENTARY NOTATION

Rex R. Michel, Contracting Officer's Representative.

17. COSATI CODES			18. SUBJECT TERMS (Continue on reverse if necessary and identify by block number)
FIELD	GROUP	SUB-GROUP	Knowledge elicitation, Knowledge engineering
			Decision making, expertise. (SDW)
			Situation assessment,

19. ABSTRACT (Continue on reverse if necessary and identify by block number)

A Critical Decision Method (CDM) has been developed for knowledge elicitation. The CDM, an extension of the critical incident technique, includes protocol analysis and memory recall tasks to study cognitive performance. A set of probes is employed to trace the development of situation assessment during critical incidents, and to determine the decision strategies used. The outputs of the method include inventories of the critical cues, graphic portrayals of the situation assessment process, and categorization of the decision strategies. Thus far, the method has been used with a variety of decisions and appears especially well suited to studying cognitive performance in naturalistic settings. It also appears valuable for addressing the highly skilled decision maker, and for eliciting the analytical and perceptual bases of proficient performance. Applications have been made for training, decision support systems, and the development and evaluation of knowledge-based systems.

20. DISTRIBUTION/AVAILABILITY OF ABSTRACT: ☐ UNCLASSIFIED/UNLIMITED ☒ SAME AS RPT. ☐ DTIC USERS	21. ABSTRACT SECURITY CLASSIFICATION: Unclassified
22a. NAME OF RESPONSIBLE INDIVIDUAL: Rex R. Michel	22b. TELEPHONE (Include Area Code): (513) 767-2691 / 22c. OFFICE SYMBOL

DD FORM 1473, 84 MAR — 83 APR edition may be used until exhausted. All other editions are obsolete.

Technical Report 799

Knowledge Elicitation of Recognition-Primed Decision Making

Gary A. Klein
Klein Associates, Inc.

Donald MacGregor
MacGregor-Bates, Inc.

ARI Field Unit at Fort Leavenworth, Kansas
Stanley M. Halpin, Chief

Systems Research Laboratory
Robin L. Keesee, Director

U.S. ARMY RESEARCH INSTITUTE FOR THE BEHAVIORAL AND SOCIAL SCIENCES

5001 Eisenhower Avenue, Alexandria, Virginia 22333-5600

Office, Deputy Chief of Staff for Personnel

Department of the Army

July 1988

Army Project Number
2P665502M770

Small Business
Innovative Research

FOREWORD

The U.S. Army Research Institute for the Behavioral and Social Sciences (ARI) is concerned with the effectiveness and user acceptability of Army decision support and training systems. A critical early step in developing quality systems is to accurately define the task(s) to be supported or trained. Using inappropriate methods to determine how job incumbents perform the task results in inaccurate task representations and ultimately in costly system design errors.

This paper describes an approach for conducting knowledge elicitation of experts that focuses on those critical components of information usage that define expertise. As such, it differs from the typical knowledge elicitation techniques that attempt to force the expert's description into the existing knowledge representation schemes of computerized expert systems and instructional design technologies. This methodology will be of value to a wide range of researchers and practitioners who are developing the next generation of training and decision support systems for the Army.

Edgar M. Johnson

EDGAR M. JOHNSON
Technical Director

EXECUTIVE SUMMARY

Requirement:

To describe the development, procedures, and products of a knowledge elicitation approach called the Critical Decision Method (CDM). The CDM is used to elicit the decision-making behavior of experts in their natural setting.

Procedure:

Review the relevant literature on decision making and knowledge elicitation and describe the theoretical basis for the CDM.

Review the findings from previous applications of the CDM and describe the development of the CDM over these applications.

Describe the procedures and products of the CDM.

Identify potential applications of the CDM in decision aid development and training.

Findings:

The CDM has been used with a variety of decision-making tasks and is especially well suited to studying cognitive performance in naturalistic, time-pressured settings.

The CDM is especially valuable for analyzing the behavior of highly skilled decision makers, and for eliciting the analytical and perceptual bases of proficient performance.

The CDM is applicable to decision-aiding problems because it identifies critical elements in task performance in terms that directly relate to the task as defined by those who perform it.

The CDM is applicable to training problems because it can identify the knowledge that sets the expert apart from others. A "data base" of expertise can then be developed as a useful resource for the direct transfer of upper-level knowledge to others.

Utilization of Findings:

The Critical Decision Method of knowledge elicitation offers a valuable alternative or adjunct to the more abstract elicitation methods associated with expert system and instructional design technologies. Because it is based on models of how experts actually make decisions and not on so-called "ideal" unnatural models, the CDM captures the nuances of expert performance that other methods typically fail to elicit. Thus, decision support and instructional system designs that incorporate CDM analysis results are more apt to be accepted and used by the target population. It is recommended that the CDM approach be used for knowledge elicitation in defining systems where expert performance definition is important, especially where the expert must perform under time pressure.

KNOWLEDGE ELICITATION OF RECOGNITION-PRIMED DECISION MAKING

CONTENTS

	Page
INTRODUCTION	1
Issues in Knowledge Elicitation	1
Knowledge-Base Development	1
REVIEW OF KNOWLEDGE ELICITATION METHODS	7
Memory Recall and Reconstruction	7
Cloze Experiments	8
Multidimensional Scaling	8
Protocol Analysis	9
Personal Construct Methods: Repertory Grid	10
Lens Model/Regression Approaches	11
DEVELOPMENT OF A KNOWLEDGE ELICITATION APPROACH FOR RECOGNITION-PRIMED DECISIONS	13
Theoretical Orientation	13
Development of the CDM	20
CDM Products	28
Summary	35
DECISION-AIDING AND TRAINING IMPLICATIONS	36
Identifying Decision-Aiding Opportunities	36
Identifying Opportunities for Improvement of Training	38
SUMMARY	41
REFERENCES	43
APPENDIX A. INTERVIEW GUIDE FOR THE URBAN FIREGROUND KNOWLEDGE ELICITATION	A-1
B. EXEMPLAR INCIDENT ACCOUNT FROM THE URBAN FIREGROUND KNOWLEDGE BASE	B-1

LIST OF TABLES

Table 1. CDM probe types and their capabilities	23
2. Incident probes for urban fireground commander knowledge base	26

Page

Table 3. Decision probes for wildland fireground commander knowledge base 27

4. Situation assessment checklist for urban fireground commander knowledge base 29

5. Situation awareness record 30

6. Model of decision point analysis for urban fireground commander knowledge base 33

7. Model of decision point analysis for wildfire commander knowledge base 34

LIST OF FIGURES

Figure 1. Model of knowledge base development process 2

2. Concurrent model of option evaluation 16

3. Serial model of option evaluation 17

4. Recognition-primed decision model 18

5. Schematic representation of advance to contact 31

KNOWLEDGE ELICITATION OF RECOGNITION-PRIMED DECISION MAKING

Issues in Knowledge Elicitation

Over the past decade, extensive effort has been applied on a number of fronts to improve the quality of human performance in decision-making tasks. The means utilized in these efforts have been extremely varied and include (a) the development and implementation of technologies to aid and support cognitive and behavioral components of human performance; (b) the design of instructional curricula to speed the training of individuals to expert-level proficiency in task performance; and (c) the design of systems to automate critical task functions that incorporate elements of human reasoning processes. The models used to represent the development process associated with each of these areas are as diverse as the areas themselves. However, a common element that exists in all efforts to improve human performance is a decomposition of human knowledge about a task into a model that identifies components in task performance that can be enhanced through training, aiding or automation.

One approach for improving the overall level of human performance in a task is to develop a thorough understanding of how a particularly proficient individual performs that task. By studying in detail the general knowledge, specific information and reasoning processes an expert uses in task performance, a model of the task can be constructed that exhibits some of the properties of the expert modeled. In addition, the model can be used to identify opportunities for improved training of non-experts and for aiding and supporting decision-making in general.

A critical question for the development of such systems concerns how knowledge bases should be constructed. In general, what theories, methodologies, techniques and representational languages should be used to model a particular decision problem? More specifically, how should the process of eliciting knowledge from experts be designed and what considerations should be incorporated in developing and selecting an elicitation technique?

Knowledge-Base Development

A knowledge base is an organized collection of information pertaining to the performance of a specific task. Knowledge bases differ widely in substance and form depending on (a) specifics of the task for which they are constructed, (b) the purpose for which they are constructed and (c) the methodologies used to construct them. Details of these differences will be discussed later as they pertain to training and decision aiding. However, the general process used in knowledge elicitation and knowledge-base development can be represented in terms of the simplified model shown in Figure 1.

This is an iterative model of knowledge base development. Iterative here means that a knowledge base is developed by applying a set of techniques to build a prototype. The prototype is refined by a repeated process of checking for completeness, consistency, and validity. This "test-operate-test" design philosophy underlies the general development process for most knowledge-based systems.

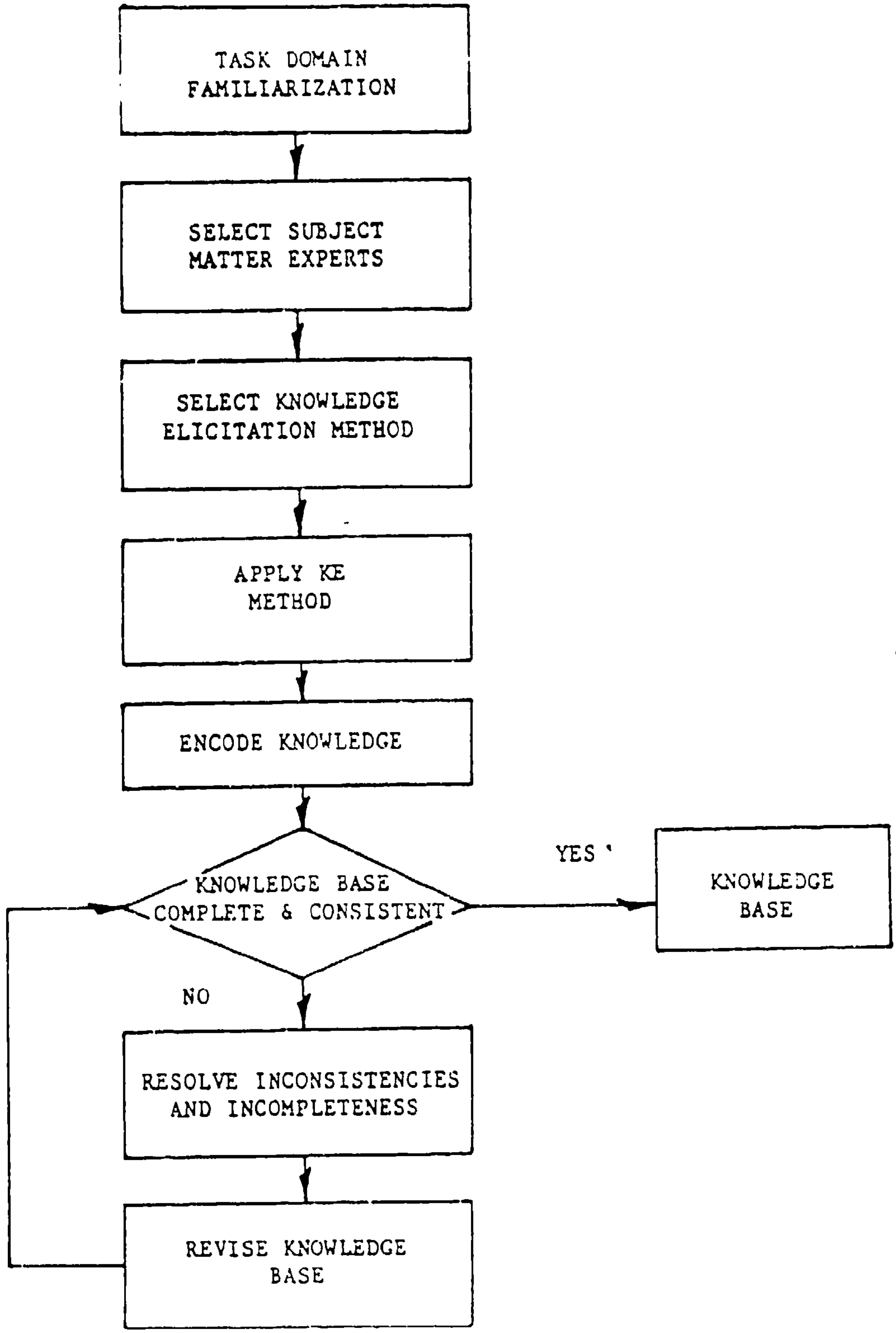

Figure 1. Model of Knowledge Base Development Process.

The process can be modeled as a number of discrete steps, each of which is associated with a set of possible methodologies. Also associated with each step are decisions and problems that the knowledge elicitor must consider as part of the elicitation effort. The model presented in Figure 1 is, of course, very general and is not intended to convey the complexity associated with modeling a specific knowledge domain. It is a synthesis of design advice gleaned from a number of resources on the knowledge-base construction process (e.g., Buchanan & Shortliffe, 1984; Hayes-Roth, Waterman, & Lenat, 1983; Waterman, 1986). However, it highlights the central issues in knowledge base development and provides a vehicle for discussing critical elements in the design of knowledge elicitation tools.

Task domain familiarization. An early, if not initial, step in knowledge elicitation is familiarization with the knowledge requirements necessary for task performance. There are different aspects of such knowledge; some of it is explicit, and some is tacit knowledge. We can distinguish between general world knowledge, specific domain knowledge of rules and procedures, specific domain knowledge of causal interactions between variables, and perceptual-motor "knowledge."

One category of knowledge is domain-independent and includes facts about the world in general. For example, rules about arithmetic might need to be brought to bear by an expert to solve problems of interest.

A second category of knowledge includes knowledge related to all facets of the task including facts, terminologies, definitions, parameters, and procedures. These facts are domain-specific in that they apply only to the domain of interest. This category of knowledge is relatively expensive to obtain since it must be elicited anew for each new domain. It is very powerful, however, since it is specifically tailored to that domain. Sources for such knowledge can include both training manuals and instructional materials, as well as direct interviews with domain experts. While the former source (written documentation) is often more readily available, its adequacy is in many cases questionable and direct elicitation of experts is called for. Elicitation techniques designed to organize and catalog domain facts are useful for this phase of the elicitation process.

The third category to consider is the mental model of the domain, including the effects of manipulating different variables.

Catalogs of domain facts, rules, and procedures alone, however, are insufficient to capture the qualitative aspects of expertise. First, expertise is difficult to define in terms of an aggregation of microscopic skills or performance elements. Just as language fluency cannot be explained in terms of phonetic rules, syntax, or vocabulary, expert performance cannot be explained by resorting to decomposed taxonomies of skills. The constituents of expertise reside in an understanding of the processes by which skilled performance evolves and in explanations of performance that emphasize tacit knowledge obtained through experience (Polany, 1966). Identifying the skills that make up expertise must begin by first viewing expertise phenomenologically in terms of the holistic relationship between the expert and the task being performed (e.g., Dreyfus, 1979; Klein, 1978). The internalization of critical task elements achieved by the expert contributes heavily to the fluidity of performance one typically observes in individuals of this category.

Second, expertise often occurs only at the boundaries of a problem domain. A large portion of the problems that experts are called upon to solve are ones for which many trained people have experience. Consider a medical diagnostic task. Many of the diagnostic problems that medical practitioners observe are quite routine in nature and may not require special expertise beyond that normally acquired by most physicians in the course of their professional practice. However, some problems are experienced only rarely and then only by individuals who specialize in solving them. Low base-rate problems lie outside the bounds of routine or "textbook" knowledge for a field. These are the problems that require a deeper understanding of the problem domain than an advanced beginner would possess. Relevant elements here include ways of characterizing or representing situations, critical information cues that are not a part of typical task descriptions, and causal factors that do not appear in rule or procedural descriptions.

Selection of subject matter experts. The selection of appropriate subject matter experts is one of the more critical elements in knowledge-base development. Two general categories of problems arise in the selection process. The first category relates to the availability of expert personnel. Obviously, one wishes that the expert under elicitation is the most experienced and possesses the most of what is defined as expertise. This is less easily achieved than might be assumed. First, knowledge elicitation is a costly and time-consuming process. Senior and, presumably, more expert people may be too busy and, therefore, unwilling to participate in a lengthy elicitation process. Junior individuals may have more time available and are, from an institutional perspective, better candidates. Second, the knowledge elicitation process itself is inherently interactive. Those who are less verbal may be disinclined to participate in a process that they find uncomfortable even though they may be the best qualified from an expertise standpoint. Third, institutions may think of expertise in terms of years of experience or rank within the organization. These criteria are not always as well related to performance as one would like. Junior individuals having better task skills may not be afforded the opportunity to demonstrate those skills in a way that would identify them as having outstanding abilities.

A more difficult problem for the selection of subject matter experts is developing a suitable operational definition of expertise. The previous section on task domain familiarization touched upon some of the problems in this area as they pertain to identifying skill components. However, the problem is broader and more pervasive. Proper definitions of expertise cannot be based on institutional or social factors alone. An ability to achieve power and position with an organization through political or personal means says little about the psychological components of expertise, except to the extent that knowledgeability plays a role in the development of organizational or group hierarchies. For example, groups of experts may order themselves according to apparent rules; one individual is more expert than another if they know what that person knows and more. One way to conceptualize an expertise dimension is to perceive the set of personnel ordered on the basis of additional knowledge. Thus, one person is more expert than a second if person B is sensitive to all the causal factors that person A knows about, but can appreciate additional factors or can recognize circumstances that change the operations of certain causal factors. Similarly, person C would be sensitive to causal factors person B might not recognize, and so on, creating a set of transitive relations. Experts themselves often try to establish such hierarchies when they compare notes and pose real and hypothetical questions to each other.

Unfortunately, social processes may do little to reveal the critical knowledge that experts ultimately demonstrate when faced with real problems to solve. Moreover, experts may not be willing to disagree with one another for professional reasons, making their public statements less useful for identifying relative levels of expertise.

Definitions of expertise based on cognitive factors are more relevant to the development of valid knowledge bases. However, specifying the cognitive factors that distinguish expert and novice performance is difficult. First, expertise is evolutionary, implying a metamorphic and developmental process rather than an incremental one. Evolutionary processes are better characterized by continuous change rather than discrete stages that can be readily defined. It does appear, however, that experts are more aware of subtle causal elements in the tasks they perform than are novices (e.g., Dreyfus, 1979; Klein, 1978) and may have a better facility for making predictions about the future state of processes they are commanding or controlling (e.g., Rouse & Morris, 1986; Klein & Peio, 1982).

Selecting a knowledge elicitation method. Selecting a knowledge elicitation method involves a number of important decisions. A useful metaphor is that of a tool box containing a number of knowledge elicitation tools that can be applied to a particular problem domain. The problem is one of deciding which tool to select.

A critical consideration in selecting a knowledge elicitation method is its cognitive compatibility. A knowledge elicitation method is cognitively compatible to the extent that the structure and process used to elicit the knowledge is consistent with how that knowledge is naturally thought of or expressed by the expert (e.g., Rouse & Morris, 1986). Compatibility can be achieved in a number of ways. One way is to approach the knowledge-modeling problem in terms the expert can understand. Calling for the expert to provide information about elements of the task that are unfamiliar or are presented in novel ways may lead to disappointing results. Adopting a normative but unnatural task model to guide the elicitation process may force upon the expert a problem representation that is not meaningful and require the making of judgments and decisions in ways that are uncustomary or wrong. For example, a fault tree representation may appear to be a useful tool for guiding the elicitation of diagnostic knowledge, but if a diagnostician is not accustomed to thinking in fault tree terms the requirements of the model may lead the expert into an unnatural task or into a misrepresentation.

A general class of methodologies for knowledge elicitation involve direct interaction with an expert in an interview type of format. The interview is given a structure provided by a theory or model of what the knowledge elicitor hopes to reveal about the contents of an expert's train of thought or the processes the expert uses in problem solving or decision making. Often referred to as process tracing, these techniques seek to provide detailed descriptions and/or models of specific aspects of task performance in terms that directly relate to the information requirements and cognitive processes associated with the task.

While such techniques have frequently been used as a vehicle for better understanding of cognition in general, they also have application as tools for modeling expert knowledge and can assist in identifying critical underlying components of expert performance. Process tracing recognizes and preserves the holistic character of decision making and avoids artificial decomposition of a problem. The knowledge elicitor using process tracing seeks to understanding how the expert solved a problem in the expert's own terms. Contextual richness is retained and incorporated in the elicited knowledge base rather than controlled for and excluded, as is done in laboratory-like elicitation tasks. This approach is particularly suited to tapping task proficiency levels that rely heavily on an individual's overall grasp of a data-rich context, such as is evidenced in expert performance.

REVIEW OF KNOWLEDGE ELICITATION METHODS

The field of knowledge elicitation draws heavily upon all facets of psychology, both theoretical and applied. Indeed, many of the tools and techniques available to the knowledge elicitor have been applied in various fields of psychological research and for purposes other than system development. Rating scales, interview techniques, judgmental modeling, and process tracing have been used as primary methodological techniques for the empirical study of cognition and behavior. This section reviews a set of specific knowledge elicitation methodologies that can be applied in a systems development context and discusses both their strengths and weaknesses. The elicitation methodologies reviewed here do not exhaust the domain of what could be applied to knowledge modeling, but are representative of techniques currently employed in knowledge-base development.

Memory Recall and Reconstruction

Perhaps the most widely utilized technique for constructing expert knowledge bases relies heavily on the ability of a knowledge elicitor to extract from the memory of an expert information in the form of facts, rules, and procedures used to make a decision or solve a problem. This is usually done by asking the expert to recall cases that are memorable for a variety of reasons. They may be cases that are representative of the typical case with which the expert deals. Or, they may be cases that are extreme or unusual, and illustrate the application of a somewhat infrequently used principle or set of facts (e.g., Waterman, 1986; Hayes-Roth, Waterman, & Lenat, 1983). For example, a knowledge engineer constructing a knowledge base for a medical diagnosis problem might first ask the medical expert to recite as many procedural details as possible about a routine case (or set of routine cases) within the diagnostic domain chosen for modeling. The procedural details would be given a formal representation according to some representation scheme (e.g., Winston, 1984) and would comprise a portion of the overall expert knowledge base dealing with cases forming the central portion of a distribution of case typicality. In addition, the knowledge elicitor would ask for a description of an unusual or infrequently occurring case, often ones with which only the expert will have had any detailed experience. Those cases probe the extremes of the expert's experiential base and provide additional procedural details for the expert model.

The validity of this approach to constructing expert knowledge bases rests in psychological theories of memory that emphasize the importance of experience as a fundamental organizing principle. One of the most important of these theories proposes that memory is organized in terms of "episodic" encoding, according to which instances of experience are remembered in terms of features such as temporal relatedness, spatial orientation, and co-occurrence (e.g, Tulving, 1972).

Though memory is a principal means by which one accesses the kinds of historical data needed for knowledge-base construction, it is important to recognize its potential flaws. Memory processes themselves are often influenced by factors outside of their contents. For example, what one retrieves from memory can be biased by temporal factors associated with storage and ease of recall. The latter has been termed an "availability" bias in the memory for events (e.g., Tversky & Kahneman, 1973, 1974), a tendency for memory to be influenced by factors such as vividness, salience, and recency.

One means of facilitating memory retrieval is to provide a framework to stimulate the process. Indeed, the elicitation setting is a major component in that framework. However, the presence of a framework does not guarantee its completeness. Some research suggests that people are typically very poor at gauging the degree to which a structure they have been given to assist them in remembering components of a problem is a complete structure, and may actually judge it to be more complete than it actually is (e.g., Fischhoff, Slovic, & Lichtenstein, 1978). Therefore, the guidance an elicitor gives to the recall process must be considered a source of potential bias in the overall picture obtained from the expert. Failure to do so may lead to unwarranted confidence in the quality of the knowledge base elicited.

Cloze Experiments

A cloze experiment is one in which an elicitor presents an individual with a passage that systematically omits a selected set of words, terms, quantities, concepts, or logical linkages. The subject is then called upon to provide the omitted terms. Though the methodology could be used as a basis for criterion-oriented testing when the correct or normative terms are known, it can also serve as a projective or knowledge elicitation tool. For example, consider the context of building an expert model for a diagnostic task. The elicitor would present the expert with a scenario (in either written or oral form) having a set of diagnostic cues and logical relationships, but omitting the diagnostic conclusion, which the expert would be asked to provide. Alternatively, the diagnostic conclusion could be given and one of the cues omitted, in which case the expert would provide the appropriate level of the cue consistent with the given cues and the diagnostic category provided. By carefully selecting a range of scenarios and systematically omitting decision-relevant terms, the elicitor can construct a model of the expert's decision processes on the basis of a series of cloze tests.

However, cloze tests comprise a process only and give no guidance on either the construction of appropriate scenarios or the t pes and ranges of terms to be elicited through selective omission. A separate methodology is needed to determine the dimensional structure of the cloze experiment (such as multidimensional scaling or multivariate factorial designs). The cloze experiment also brings to the elicitation context a risk of calling upon the expert to reason in a format that bears an unnatural resemblance to the way they routinely operate. Though this criticism can be applied to elicitation formats in general, the problem may be particularly acute for the cloze test when important qualifying details are omitted from scenario passages in interests of structural simplicity or design considerations.

Multidimensional Scaling

Multidimensional scaling is an empirically-based methodology used to systematize and represent a domain of objects in terms of a set of dimensions (e.g., Schiffman, Reynolds, & Young, 1981). The method can be used to map the conceptual elements of a knowledge domain and can serve as a framework for understanding how they are related. It is usually applied by first generating a list of conceptual elements for a particular task. These elements are usually the concepts associated with the task or are linguistic elements used to describe the task. For example, a fire fighting task might contain conceptual elements such as "cooling down" and "ventilation." The set of elements is presented to the expert as a series of pair-wise comparisons to be judged in terms of their similarity (e.g., "How similar are cooling down and ventilation?") The resulting judgments are analyzed using a multidimensional scaling procedure to determine clusterings of the elements. The analysis provides a mapping of the knowledge domain in terms of conceptual similarities between domain elements. Tightly clustered elements suggest cognitive linkages that can be exploited using more detailed elicitation techniques or as guidance for structuring an interview process. For an excellent example of the technique as applied to the problem of modeling aircraft-piloting expertise, see Schvaneveldt, Durso, Goldsmith, Breen, Cook, Tucker, and De Maio (1986).

Principally, multidimensional scaling has been used as an analytic technique for mapping a multidimensional space containing a set of complex elements and can be used to form the basis for clustering common elements and identifying outlying or unusual ones. Though it is a potentially powerful analytical technique, it is principally a mathematical procedure, and has no cognitive content apart from the interpretation given to its results by a knowledge elicitor. The user must define the set of elements a priori, either by (a) resort to an underlying theory or model of the task, (b) eliciting them directly from an individual (or individuals), or (c) intuition and guesswork. While the technique can yield a useful map, it cannot define the elements that appear on the map. That definition requires a theory of the task obtained from a separate elicitation method. To its advantage, however, it can provide a reasonable initial theory of the structure of a problem domain and may serve as a useful process in guiding and structuring other aspects of the knowledge elicitation process.

Protocol Analysis

One of oldest methods of study available to psychologists involves the documentation and subsequent analysis of protocols obtained by asking an individual to think aloud about how he or she would solve a given problem (e.g., Ericsson & Simon, 1980, 1984). Systematic coding and analysis of those protocols provide the basis for a model of the individual's reasoning. Presumably, the elicitation of protocols provides a window into the individual's thinking in a form that leaves the reasoning process relatively unperturbed.

Though protocol analysis has been used to study cognitive functioning in a number of important content areas (e.g., Svenson, 1979), its application is not without controversy and qualification (e.g., Hayes, 1982). The relative familiarity of the cognitive task may bear on the quality of the protocols obtained. In general, unfamiliar tasks have the advantage that they may generate protocols providing information about some of the cognitive structuring of a problem and the creation of a solution process. Familiar tasks tend to be associated with readily available cognitive strategies for solving them. These strategies may be more stable than in unfamiliar tasks, but have the disadvantage of becoming highly automatized, rapid, and less accessible to studies of verbal protocols. Thus, familiar problems may induce a mode of automatic responding that may make it difficult for subjects under study to give an in-depth explanation of their reasoning. In contrast, unfamiliar problem areas may prompt a more deliberative thought process in a less automatic mode, thereby creating a richer protocol for analysis (e.g., Svenson, 1986).

A second feature of protocol analysis is its emphasis on promoting verbalization of otherwise nonverbal events. People are typically unaccustomed to giving verbal expression to cognitive events and may seriously distort their reasoning processes in attempting to apply language concepts to them. Moreover, the context provided by a knowledge elicitor may encourage the elicited individuals to elaborate on their reasoning processes in ways that are unrelated to their true content (e.g., asking "why" type questions), thereby suggesting to the elicitor that they know more about their cognitive activity than they actually do (e.g., Nisbett & Wilson, 1977).

A third difficulty in applying protocol analysis arises from the need to settle upon a theory or methodology for interpreting protocol responses. Though protocol analysis provides in principle a rich source of data, it provides virtually no guidance on how that data should be organized and interpreted. The problem becomes more acute when the intended use of the data is for system development purposes, in which a specific knowledge representation scheme is required. Protocol analysis conducted without knowledge representation guidance risks having limited interpretability. However, excessive structuring of the protocol elicitation process by a knowledge representation framework risks losing the flexibility and naturalness of the of the approach. Protocol analysis can be most fruitfully applied when the knowledge elicitor has (a) a general knowledge representation framework in mind, and (b) an overall model of the relationships between key knowledge components, such as might be obtained by applying multidimensional scaling.

Personal Construct Methods: Repertory Grid

Personal construct theory has been used as a foundation for a method of modeling human expertise (e.g., Boose, 1985). Its methods employ some variant of the repertory grid technique. The repertory grid is an instrument designed to elicit construct relationships (e.g., like-dislike, important-unimportant, chronic-acute) between a number of elements in a specified domain, such as people one knows, attitudes one has, or cases one has diagnosed. The grid itself is comprised of three components: (a) elements, which define the material upon which the grid is based; (b) constructs, which are the terms that an individual uses to group and differentiate between elements; and (c) linking mechanisms, which show how each element is assessed on each construct. A critical part of constructing the grid is the selection of elements, for it is this initial selection that defines the relevant judgmental domain. For practical purposes, it is important that the elements be as specific as possible, or else the basis for judgment will be too vague. Important design dimensions include numerosity (too many elements make the task cumbersome, too few make the space less relevant), homogeneity, and representativeness of coverage.

While the elements of the grid are the objects of one's thoughts, the constructs are the qualities one attributes to those objects. The traditional method for generating constructs is to elicit them from triads: triples of elements are presented to an individual who is to respond with the way that two of the elements are alike and in what way the third element is different from the two. Though this procedure produces polar-opposite dimensions, it is important to note that the polar distinction is somewhat forced by the elicitation process and can produce an apparent orthogonal dimension structure when one is absent. Furthermore, the process can produce logical opposites rather than constructs that are logical in meaning. For example, the logical opposite of "ambiguous" is "unambiguous." However, a subject may think of the real opposite of "ambiguous" as "things I understand;" clearly the two polar dimension sets are not equivalent.

Choosing the triads is a critical aspect of grid development. When the number of elements is fairly small, all elements can appear in triad sets through a modified paired-comparisons design. However, as the element set becomes large, the set of triples grows rapidly and a criterion must be established for choosing what elements get presented with what other elements. An alternative method is to present a few of the constructs in a triad approach and use the resulting constructs as a stimulus set for directly eliciting other constructs (e.g., Easterby-Smith, 1980).

Once elements have been defined and constructs elicited, the next step is to link the elements through the constructs. This is typically done by having the individual rate each element on the set of bipolar construct scales. The resulting judgments allow the elicitor to cluster and differentiate the elements in terms of the construct framework. Though several methods exist for analyzing repertory grid data (e.g., Klein & Klein, 1978; Shaw & Thomas, 1978), interpretation is more clinical art than methodological technique, and the knowledge elicitor must have a general theory of the knowledge domain to guide that interpretation. That theory, however, could be developed from an initial application of a personal construct methodology, could be based on another of the elicitation methodologies described in this section, or could be based on a particular knowledge representation scheme.

Lens Model/Regression Approaches

An important category of tasks are those that call upon an individual to render a judgment or decision on the basis of a number of pieces of information. A paradigmatic approach for making explicit the processes involved in multiple cue judgment is the lens model (e.g., Hursch, Hammond, & Hursch, 1964; Dudycha & Naylor, 1966). Briefly, the lens model draws a distinction between an object that is represented by a collection of information cues, and the psychological representation of that object in terms of a judgmental policy based on those cues. Like an optical lens, the environment is portrayed in terms of a set of information components with predictive importance expressed in terms of validity coefficients. Information cues are used by the individual to form a judgmental impression of the object in the environment. The policy by which the cues are weighted and an impression is formed can be modeled as a regression equation. The equation predicts the individual's judgment of the object from a linear combination of cues and cue weights derived empirically by varying (usually systematically) the information features in the environment and observing the judgmental response. Decomposition of the regression equation permits a determination of the relative emphasis the individual places upon the various information components.

A principal advantage of the lens model approach is that it elicits a judgmental response from an individual in a manner that tends to preserve many of the natural features of the environment in which the response is typically made. No attempt is made to probe directly the individual's reasoning processes. Rather, the emphasis is on modeling response policy through selectively varying environmental features. However, while a model obtained this way may bear good fidelity in a predictive sense to the behavior of the individual, there is no guarantee that the parameters within the model are isomorphic to the underlying cognitive processes that produced the response. Despite this limitation, lens modeling is a valuable tool for assessing the relative weight an individual gives to information in a judgmental task and can be used as a methodology for corroborating the results of more cognitively intrusive elicitation techniques.

DEVELOPMENT OF A KNOWLEDGE ELICITATION APPROACH FOR RECOGNITION-PRIMED DECISIONS

Theoretical Orientation

Up to this point in the report we have been discussing general issues in knowledge elicitation. Now it is time to focus more directly on the Critical Decision method (CDM) that we have developed. The CDM is a form of protocol analysis and process tracing. It differs from standard protocol analysis because it involves interviewing with subjects about prior events, rather than asking subjects for think-aloud reports of ongoing problem solving. Thus, the CDM gives up the important control found in most protocol analysis studies. What it gains is the opportunity to probe about critical incidents. Flanagan (1954) has developed a Critical Incident method for job analysis, and the CDM is in some ways an extension of this method, adapting it for the knowledge elicitation of naturalistic decision making.

The CDM is also similar to memory recall and reconstruction methods since it relies on retrospective accounts of incidents. It differs in the emphasis on perceptual cues and patterns, the sensitivity to tacit knowledge rather than rules, and the theoretical orientation on situational awareness and recognition-primed decisions.

The rationale behind the CDM can be visualized in the following way. Much of what an expert knows has become fairly routine. We can probe this knowledge using standard task analytic methods but it is difficult to learn much because the knowledge has become automatic and unconscious. We can also probe about novel cases, and this is what we do in many laboratory settings when we use unfamiliar tasks. Such probing can tell us about reasoning abilities independent of content. But it will not tell us much about content expertise since by definition the subject is a novice.

In-between these two strategies is a third approach -- to probe about non-routine incidents that required expertise, and were carried out effectively only because of that expertise. This is the approach taken by the CDM. We have used it to study expertise in a variety of settings.

ROUTINE CASES	----- TASK ANALYSIS
NON-ROUTINE CASES	----- CDM
NOVEL CASES	----- BASIC RESEARCH

The CDM is an interview method for eliciting some of the bases for expert performance. The method is focused on the tacit knowledge of the expert in addition to the explicit knowledge. The emphasis is on non-routine events, usually recent ones that occurred in operational settings. (That is why we could not use protocol analyses, since operational non-routine incidents cannot be scheduled, and subjects are usually unwilling to provide think-aloud protocols.)

In other words, we feel that it is at the boundary between routine and novel where expertise is stretched far enough to become visible.

The method is grounded in recent theoretical analyses of naturalistic decision making by experts. The CDM has evolved from a different theoretical perspective on decision making, one that is more attuned to naturalistic decision making, especially under time stress. The next section will describe this theoretical perspective, and the following sections will describe the CDM itself.

Behavioral Decision Theory. Decision making vis-a-vis the decision sciences has emphasized the application of normatively appropriate decision models as a basis for decision making. The most widely advocated of these models is decision analysis (e.g., Raiffa, 1968; Keeney & Raiffa, 1976). A decision maker acting in accord with the tenets of decision analysis will:

a. enumerate all possible courses of action;

b. identify the possible outcomes associated with each course of action;

c. place values on each outcome;

d. for each course of action assess the probability that the associated outcomes would be experienced;

e. multiply each outcome by its associated probability to obtain an expected value or utility;

f. sum the expected values for each course of action;

g. choose the course of action having the largest expected value.

While normative decision theory has focused on what people should do in decision-making situations, behavioral decision theory has studied what they actually do. One thrust of behavioral decision theory has been to understand the degree to which people are capable of providing the kinds of inputs required of decision analysis; another has been to understand the mental rules or strategies they naturally use in making decisions.

The vast literature on these two points will not be reviewed here. The general finding has been that people are very poor decision analysts at best. They are typically not good at generating complete sets of options (e.g., Gettys & Fisher, 1979; Pitz, Sachs, & Heerboth, 1980), outcome evaluations are subject to strong context effects and intransitivities (e.g., Lichtenstein &

Slovic, 1971; Fischhoff, Slovic, & Lichtenstein, 1980), uncertainty assessments are poorly calibrated and exhibit an insensitivity to the degree of actual knowledge (e.g., Lichtenstein, Fischhoff, & Phillips, 1982; Fischhoff & MacGregor, 1982), and their choices do not necessarily abide by the principles of utility maximization (e.g., Fischhoff, Goitein, & Shapira, 1982; Simon, 1955).

A second thrust of research in behavioral decision theory has been to identify informal rules or heuristics people use in solving routine cognitive problems and making decisions. The majority of these heuristics are very useful for dealing with most problems people encounter in everyday life, but can lead to predictable errors and judgmental biases in some circumstances. A commonly used heuristic for judging the uncertainty of a proposed event is availability (Tversky & Kahneman, 1973) whereby an individual relies upon the ease with which relevant instances can be recalled from memory as a basis for judgment. However, ease of recall can be influenced by factors unrelated to the actual frequency with which they are experienced, such as vividness, salience, and recency. Thus, events that are particularly memorable because of features unrelated to their frequency of occurrence can exert a biasing effect on judgments of relative likelihood. Interview techniques designed to elicit likelihood judgments from an individual regarding elements of an event need to be sensitive to how the memorability of the event may influence judgment.

A second important judgmental heuristic is representativeness. Judgments based on representativeness focus on the degree to which a particular case or instance shares features similar to a prototype, class, or stereotype (Tversky & Kahneman, 1974). Thus, likelihood judgments about the occurrence of an event will be greater the more features the event has in common with the class from which it comes. Judgments by representativeness are made according to the associative strength present between an event class and a particular event. A biasing effect on likelihood judgments occurs because the heuristic leads to an insensitivity to the distributional properties of samples. Small samples are regarded as more predictive than is appropriate (Tversky & Kahneman, 1971) and event patterns are judged as meaningful on the basis of appearance alone, without resort to the underlying uncertain process which generated them (Kahneman & Tversky, 1972). For example, a battle commander might be fooled into reacting wrongly because the enemy presents to him a subset of activities that are like those that appeared previously when the enemy took certain actions. The commander ignores the other actions these activities might also indicate; basing his reaction solely upon a single previous event and a limited subset of the possible indications.

Both the availability and representativeness heuristics are based on the natural psychological processes people use in responding to the tasks and problems posed to them. In complex contexts such as knowledge elicitation of experts, their biasing effects are much more difficult to gauge than in well-constructed, simplified laboratory tasks. Nonetheless, they play an important role in the elicitation process by biasing the events an expert chooses to describe.

<u>Recognition-Primed Decision Making</u>. While behavioral decision theory has emphasized biases in the understanding of human decision making, a different strategy is to try to understand and model the natural decision processes

themselves. Many of the biases could be specific to laboratory paradigms, well-defined tasks, and the use of naive subjects. What are the decision strategies used by experts in field settings?

A major finding of previous research in naturalistic decision making is that experts rarely report considering more than one option (e.g., Klein, Calderwood, & Clinton-Cirocco, 1985). Instead, their ability to handle decision points appears to depend on their skill at recognizing situations as typical, or as instances of general prototypes developed through experience. These prototypes provide an understanding of the causal dynamics associated with a decision problem, suggest promising courses of action, and generate expectations.

One way of contrasting this conceptualization of decision processes with more a normative approach to decision making is in terms of serial and concurrent models. Figure 2 portrays a standard decision analytic model: options are enumerated down the side and a set of relevant evaluation dimensions across the top. According to this model, the decision maker considers several options at the same time. This can be done by performing pair-wise comparisons of options in terms of the evaluation criteria, making concurrent, conscious judgments and relative evaluations of the strengths and weaknesses of each option. Options are then ordered according to their performance on the evaluation criteria and the option selected for implementation is the one achieving the highest multi-attribute evaluation score. Though decision models of this general type are widely prescribed as a basis for making decisions in many contexts (e.g., Raiffa, 1968), their validity as descriptors of human decision making is extremely weak. In time-sensitive contexts, for example, the tempo of decision making is frequently much too fast for the performance of an all-inclusive generation and evaluation procedure (e.g., Klein et. al., 1985).

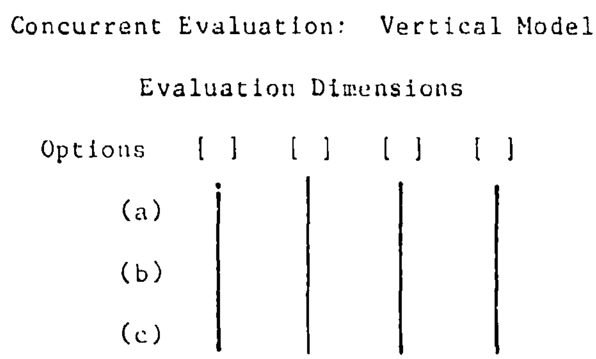

Figure 2. Concurrent model of option evaluation.

In contrast, Figure 3 shows a serial option evaluation model. Here, an option is generated, tested for feasibility, and then either implemented or rejected. If it is rejected, a second option is considered, and so forth, until a suitable option is found. All options are not necessarily evaluated on all dimensions or on the same dimensions. This may be described as a horizontal model of decision making: although one or more options may be considered, only one option is examined at a time. The description of decision making provided by this model is a more consistent process description of what might occur in most applied, time-sensitive contexts. A description of the categories of decision making processes is found in Svenson, 1979.

Serial Evaluation: Horizontal Model

Evaluation Dimensions

Options [ ] [ ] [ ] [ ]

(a) ______

(b) __________

(c) ______________

Figure 3. Serial model of option evaluation.

The process elements of the model portrayed in Figure 3 are developed in greater detail in Figure 4. This model we term Recognition-Primed Decision Making, (RPD).

Figure 4 shows three aspects of the RPD model, varying in complexity. In the simplest case (A-1), the automatic RPD, the decision maker (DM) uses knowledge and cues to recognize a situation as familiar or typical. This "automatic" recognition includes with it a recognition of what goals can be achieved, what cues to monitor, and other types of expectations. The recognition of typicality also leads to the recognition of the typical reaction, and this is implemented. An example would be a fireground commander realizing that a simple fire in a laundry chute has gotten out of hand and has spread through the top floor of an apartment building, so that it is time to call in a second alarm.

In a more complex case (A-2), the verified RPD, the decision maker again recognizes the typicality of the situation but has the time to evaluate the option, perhaps to imagine it being carried out. No other options are considered, even though they exist in an "action queue" of options varying in their availability, but the implementation of the typical option is not automatic. An example is a commander of a rescue team who knows what he wants to do but still plays the images out in his mind once or twice before issuing the order.

The most complex case of RPD involves serial consideration of options (A-3). Here the evaluation is more serious. The favored option may be implemented, or it may be modified to fit the needs of the current situation. In some cases it is rejected and the next most typical option is selected from the action queue. An example is a commander trying to rescue an unconscious woman from an overpass. He considers the most typical harness rescue, realizes that it will not work, considers another standard piece of rescue equipment as a second option, rejects that one, comes up with a third option, a ladder belt, thinks it through, and then implements it. A series of options were considered, but there was never a comparison of the merits of one option vs. another.

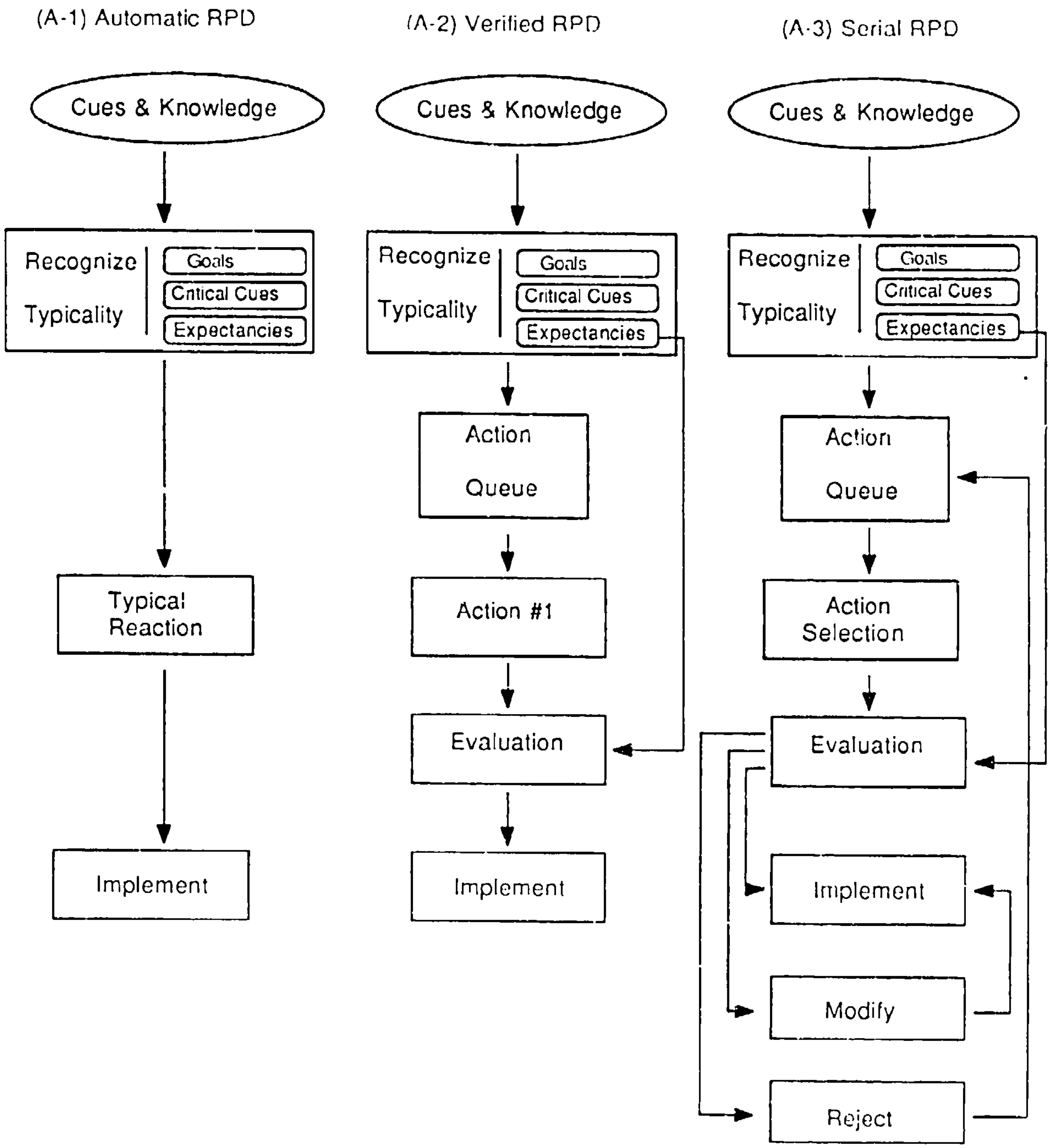

Figure 4. Recognition-primed decision model.

Actually, there are some additional components to the model to reflect how the decision maker can verify or go through serial search for the recognition of the situation, as well as the selection of the response.

What unifies all of these levels is that the decision making begins with a pattern recognition (i.e., recognize typicality in Figure 4). The decision maker uses all of the experience (i.e., cues and knowledge) gleaned from years of practice to view an event as typical in some way that resembles Rosch and Mervis' (1973) judgment of prototypicality. The recognition makes it obvious what can be accomplished, what dangers exist, what critical cues must be monitored, what expectations to form. The recognition also carries with it a realization of a typical set of reactions, and the most typical is considered first. This is quite efficient, since in most cases the most typical reaction will be the one called for. No other work will be necessary, and that reaction can be implemented. Sometimes, the decision maker may attempt to verify this option by evaluating the conditions and the plausibility of successfully carrying it out. And sometimes the decision maker may realize that it won't work, that it needs to be modified. In the most complex case, the option will be rejected as unworkable and another, somewhat less typical reaction will be considered. But at no time will the decision maker attempt to generate more than one option at a time, or to evaluate options by contrasting strengths and weaknesses.

How does this model differ from the ones postulated by standard behavioral decision theory? Those models, whether linked to Decision Analysis, Multi-Attribute Utility Theory, or Elimination-by-Aspects, all focus on the options and ignore the predecision processes. They are prescriptive models for applying the most powerful methods for evaluating a set of options. They are not designed to handle strategies such as growing a better option by modifying one that is almost adequate.

One of the most serious drawbacks of analytical decision strategies is that they take time to carry out. Studies have shown that analytical decision strategies are not effective when there is less than one minute to respond (Howell, 1984; Zakay & Wooler, 1984; Rouse, 1978). And these studies were performed with tasks that were well-defined and clearly amenable to analytical decision strategies. What would happen if the researchers used tasks that included ambiguous and missing data? How many decision makers, feeling the pressure of time, would begin an analytical process knowing that it might not generate a useful answer at the end?

For these reasons it is important to supplement the standard decision paradigms and decision models growing out of limited laboratory paradigms. The research is typically done with artificial tasks, with limited contextual nuances, using naive decision makers, ample time for decisions, and no personal consequences for making poor decisions. The paradigms do not generalize well to natural environments where there is usually time pressure, limited resources, dynamic and ambiguous goals. These paradigms have resulted in decision models that are extremely powerful but rarely applicable to command-and-control situations.

Because the RPD model assumes that an acceptable course of action may be chosen without conscious generation and evaluation of alternatives, the emphasis in this model is on what we have called situational awareness. This view is similar to the ability of coup d'oeil described by von Clausewitz (1943), the skill of making a quick assessment of a situation and its requirements. The RPD model is also consistent with theories of expertise which stress the importance of perceptual/recognitional abilities over analytical power (de Groot, 1978; Dreyfus, 1979; Klein & Calderwood, 1986; Klein & Peio, 1982).

For a time-limited task, concurrent evaluation may be impossible in many cases. Even reducing the number of options and dimensions still places a heavy load on the decision maker. Descriptively, the RPD model makes available to the decision maker a course of action at every point. The decision maker begins with an initial option, and if a response is called for, this option is executed. If there is time for some evaluation, it will be examined, accepted, improved, or rejected for a second option which then becomes primed for implementation.

The RPD model may be useful in explaining some aspects of intuitive decision making (Hammond, Hamm, Grassia, & Pearson, 1984); where experts use recognitional and perceptual matching processes, it is understandable that they would find it difficult to articulate the basis for their decisions. The RPD model is also consistent with a general approach to the study of decision making that emphasizes the actual mental heuristics people use in judgment and choice. An earlier section on Behavioral Decision Theory outlined this approach in greater detail. Unlike "optimal" decision models that are primarily prescriptive, the RPD model is essentially descriptive. It is an account of a process that allows people to proceed reasonably in many decision-making situations but does not insure optimal results.

Development of the CDM

Overview. The CDM was developed as part of a naturalistic study of decision making in the context of urban and wildland fireground command. Complete details of the studies can be seen in Klein et al. (1985) and Klein et al. (1987). Highly experienced command personnel in these two domains were studied to determine the critical elements associated with expert task performance. A semi-structured interview technique was developed as an initial prototype for the knowledge elicitation methodology, loosely based on Flanagan's (1954) critical incident method.

An Interview Guide was developed that took into consideration two distinct aspects of problem structuring. First, the interview methodology was designed to be relatively unstructured and free from interviewer bias. This was intended to permit the details of decision making to emerge with the fire ground commander's own perspective and emphasis intact. Second, the interview was sufficiently structured to avoid simply collecting unrelated narratives of isolated experiences. The perspective required was one that directed the expert to focus on those elements of an incident which most affected decision making, and to structure responses in ways that could be summarized along a specified set of dimensions.

The solution to these two goals involved asking the officer to describe an incident completely before any further elicitation was done. This procedure appeared to be successful in establishing the knowledge elicitor as a listener (rather than an interrogator), and increased overall cooperation. After the incident had been related, it was explored in depth to obtain a timeline of the event. The timeline focused on representing the actual sequencing and duration of events, as well as the information and cues available at each decision point. This technique clarified incident events and resolved questions and inconsistencies in the developing knowledge base (see Figure 1 earlier). In addition, the technique reactivated memory for much of the contextual knowledge of the incident from more than a single time perspective, an approach which has demonstrated utility in obtaining accurate eyewitness testimony (Geiselman, Fisher, MacKinnon, & Holland, 1985).

The first step in the analysis was to reconstruct the narrative of the incident, attempting to capture in as rich a detail as possible the incident scene from the point of view of the expert. Notes and timelines were checked against tapes of the interview sessions and interview notes as a guideline for the quality of incident selection and completeness.

Using the timeline and incident account, each incident was then structured into the decision format which forms the basis of the analysis. A decision point was generally defined as a point in time when alternative decisions or courses of action could have been chosen or taken. The analysis of decision points documented the nature and chronology of the expert's situational assessment and each decision point obtained from the Incident Account. This analysis was largely inferential and was accomplished according to a specified coding scheme outlined in detail below.

We have used the CDM in a variety of different settings: data analysts (Klein, Maher, Zakay, & Kessel, 1987), fireground commanders (Klein, Calderwood, & Clinton-Cirocco, 1986; Calderwood, Crandall, & Klein, 1987), design engineers (Klein & Brezovic, 1986), tank platoon leaders (Brezovic, Klein, & Thordsen, 1987), wildland firefighters (Taynor, Klein, & Thordsen, 1987), paramedics (Crandall & Klein, 1987), command-and-control officers (Thordsen, Brezovic, & Klein, 1987), MIS specialistics (Crandall et al., 1988). We have not run the CDM the same way in every setting. The time availability for interviews affects the number of questions we can ask, and the needs of sponsors determines the probes we use.

Therefore, we will first describe the basic CDM strategy, and then go into details about options and criteria for using different probes.

In its most stripped-down form, the CDM is an interview method consisting of the following stages: identifying a critical incident to probe, obtaining a brief narrative description of the incident, probing about shifts in situational assessment during the incident, probing about decision points during the incident. In almost all the applications we have found it useful and often necessary to prepare a timeline of key events (cues and actions) during the incident. We have almost always found it valuable and necessary to allow the interviewee to draw a map and to describe events in the context of that map.

The identification of a critical incident depends on the needs of the study. We look for recent incidents, usually within the previous 3 months. However we have found that memory for critical incidents appears quite good, even after a year or more. These are vivid incidents, sometimes with lives at stake. They have been thought about and discussed in the intervening time. We look for incidents where experience was important, where someone with less experience would have possibly been unable to be effective. We also look for errors, since these are good sources of data about the uses of knowledge.

The brief narrative seems important to gain an overall perspective on the event, and to improve rapport with the interviewee. We try not to interrupt, but we sometimes ask for clarifications as we take notes.

The probes for situation assessment focus on the cues that were vital for the interviewee's understanding of goals, options, cues to monitor, dynamics, expectancies. We concentrate on new cues and items of information that led to a shift in situational assessment (e.g., a new goal was substituted for a previous one) or led to elaborations of goals into more specific sub-goals. We try to identify the critical cues during the incident, and the causal factors that affected strategy and expectations.

The probes for decision points focus on places during the incident where there were alternative actions possible. These may have been alternatives that the decision maker considered at the time, or alternatives that other decision makers might have selected, especially if they were less experienced. We probe the advantages and disadvantages of each option, and some of the reasons for their not being selected. Sometimes we subdivide this into probes about the interviewee's understanding of the situation, compared to alternative hypotheses, and probes about options.

Table 1 presents the types of probes we have used thus far, along with a description of the information obtained from each. We can describe different forms of knowledge as follows:

(a) structural knowledge about scripts and rules,
(b) perceptual discriminations and perceptual pattern recognition,
(c) conceptual patterns and discriminations including causal factors and their implications,
(d) analogues and metaphors,
(e) judgments of prototypicality and familiarity.

These are not intended as a theory of knowledge, and we will not want to argue that they are distinct categories. Their function is to help us to explain the types of information we get from each of the probe types in Table 1.

Table 1

CDM Probe Types and Their Capabilities

CDM Probes	(a) Structure	(b) Perceptual	(c) Conceptual	(d) Analogues	(e) Prototypes
	Forms of Knowledge				
1. DP options			X		
2. Cues		X	X		
3. Causal Factors		X	X		
4. Goal Shifts					X
5. Analogues				X	
6. Errors	X	X	X		X
7. Hypotheticals	X		X		
8. Desired data			X		
9. (Task Analysis)	X				

The probe types are as follows.

1. DP options refers to the decision points for which there were different ways to understand the situation and different options possible. For example, in studying design engineers a decision point was the refresh rate used by the computer driving a display. In the firefighter study, a decision point was whether to make a rescue using a Kingsley harness, a Howd strap, or a ladder belt.

2. Cues refers to the primarily sensory data that affected the understanding of the situation. In the firefighter study, one incident addressed the color of smoke coming out of a building, as a cue to the heat of the fire. In the tank platoon leader study, the visible terrain was often critical as a cue for anticipating the actions of the opposing force.

3. Causal factors were the dynamics affecting what was expected, what could be accomplished, etc. In the firefighter study, the composition of the building was usually a factor affecting the speed of the spread of the fire. The method of roof construction affected whether firefighters could be sent onto the roof. In the wildland firefighting study, knowledge about the temporary disruption of communications affected the path selected by a leader who needed to make sure he knew where the fire was heading.

4. Goal shifts were radical changes in situational assessment. In the firefighting study, this might be a realization that the fire was too advanced to be put out, and what was needed was to immediately evacuate the residents of the building.

5. Analogues were previous incidents that were similar to aspects of the current incident, and suggested causal factors to monitor, goals to select, or options to implement. In the firefighter study the recollection of how a billboard crashed into the street when its supports were burned away influenced a commander to order the crowds moved back.

6. Errors were admissions of poor situational assessment and/or option selection. In the tank study, the ineffective establishment of an overview position reflected the inability of the cadet officer to sense what the other tank officers could see, and what the enemy could see.

7. Hypotheticals were probes about sequences that did not occur but might have. One of the more useful hypotheticals was employed in the wildland firefighting study about what might have happened if an action had not been taken. This showed the expectancies and the reasons that motivated the action.

8. Desired data were probes about information that would have been desired to aid in the decision process. In the command-and-control study we asked about missing information and were told about the value of more specific data on enemy location, and also about the location of friendly forces.

9. Task Analysis. This is not a CDM probe, but is sometimes used as an adjunct to the CDM, to define the structure of the task. In the data analyst study we established the overall procedures and objectives of the task before presenting information about the expertise needed.

As can be seen in Table 1, the different types of probes obtain different types of information about the knowledge used by experts in handling non-routine events. How does this work in practice? The best way to answer this question is by reviewing previous CDM efforts.

In one study of paramedics, we needed to probe about perceptual discriminations in order to determine how they were able to recognize heart attack victims. We needed to probe conceptual knowledge to understand how they interpreted these perceptual data. We needed to define their judgments of prototypicality to see how they were distinguishing different types of victims.

In a study of MIS personnel there was no interest in perceptual discriminations. Rather, the emphasis was on conceptual patterns and discriminations, and prototypes. The same was true for the study of data analysts. It was also true in a study evaluating the performance of Army and Air Force loadmasters, where perceptual cues were of little relevance.

In developing an expert system for the Air Force, we needed to elicit knowledge about the survivability/vulnerability of structures under the stress of blast waves. Here, there was a need to probe about perceptual cues in data tracings, along with conceptual knowledge and prototypes.

In our basic research experiments on firefighters, all the forms of knowledge described in Table 1 are relevant.

There is one additional type of probe that we have not mentioned in Table 1, and that is about images. Images are not a form of knowledge, and yet they seem central to the way experts evaluate options they want to implement, and we try to probe them when they come up. We have seen it in virtually all the studies we have performed. An example is the firefighter who ran through several repetitions of an image of how an unconscious car-accident victim was going to be supported prior to initiating the rescue.

Interview Guide. The purpose of the interview guide is to structure the overall elicitation approach and to provide a format for presenting specific probes. The general structural elements of the guide are discussed below, followed by an example of the interview guide format as applied to a specific decision domain. We are using a study of urban firefighters (Calderwood et al., 1987) as the context for the following discussion.

1. Introduction. This section focused on introducing the expert to the elicitation process. It gave the expert a reason for the study, provided a rationale for the process and served to establish a rapport between the expert and the interviewer. Background data concerning the expert were collected during this phase of the elicitation, and included (but were not limited to) demographics (e.g., age, sex), command experience (e.g., rank, years of experience, positions held), and training history.

2. Incident selection. This segment of the elicitation focused on the selection of an incident for discussion. Guidelines were provided to help the expert select an appropriate incident. A preliminary criterion was an incident that present a particular command or decision-making challenge. The expert was assisted in this process by examples of possible command or decision-making challenges including (but not limited to) unusual risk to life, non-standard operating procedures, and mistakes or errors in judgment.

3. Situation assessment. This aspect of the elicitation focused on obtaining a description of both (a) the history of the incident and details leading to the situation being described, and (b) the processes used by the expert in making an assessment of the situation. The RPD model posits situation assessment as an initial stage in recognitional decision making. This probe was a carefully constructed step to elicit as complete a picture as possible of the expert's conscious assessment of the decision situation.

4. Goal identification. This aspect of the elicitation focused on identifying the particular goal or command objective the expert was attempting to achieve. Goal elicitation was done by direct questioning of what the expert intended to accomplish, subsidiary goals, and the decision or command objectives. Primary inputs to the situation assessment stage of the RPD model include objectives and goals. Though evaluation of goals is not a central focus of the model, they play an important role in structuring situation assessment.

5. Deliberation of decision making. This aspect of the elicitation focused on identifying the contents of the expert's deliberations during the incident. Direct questioning was used to elicit such elements as (a) options that were considered, (b) information that was required, (c) rules used to select options, (d) processes used to assist in deliberation (e.g., mental

imagery), (e) length of deliberation, and (f) conscious assessment of deliberation. Of all the probe stages in the interview, this was the most detailed. Its contents were used to understand both the situation assessment and evaluation stages of the RPD model when possible and to obtain information on analytical decision making in which options were clearly enumerated and evaluated. This material was primarily oriented towards basic research to help us evolve a better RPD model. For applied studies there would be no need to probe about conscious deliberation. The following discussion retains many of these basic research probes.

An Application of the Interview Guide. The interview guide was applied to the study of command expertise in groups of both urban and wildland fireground commanders. The fireground command task is one that requires considerable years of training and experience to perform and calls for high-level strategic and tactical decision-making abilities.

The purpose of the elicitation was to obtain a knowledge base of the critical decision-making skills used in performing this command task. An interview guide was prepared based on the general outline discussed above. A complete version of the guide for the urban fireground command study can be seen in Appendix A.

The guide first introduced the elicitation study and then elicited relevant background information about each expert. This was followed by a description of the type of incident the knowledge elicitor wanted to probe. For example, the urban fireground officer was asked "to relate all the events from the time the alarm was received, focusing on his commands and critical decisions." Specific probes were utilized to prompt the elicitation process. These are shown in Table 2 for the urban fireground case and Table 3 for the wildland fireground case.

Table 2

Incident Probes for Urban Fireground Commander Knowledge Base

Probe Type	Probe Content
Analogy	Pick the most similar/helpful case. Describe the differences.
Scenario	Does the incident fit a standard scenario that you have ever seen or been trained for?
Number of incidents	How many incidents like this have you been involved in overall? At a command level?

Two categories of probes were utilized: incident probes and decision probes. Incident probes (Table 2) were designed to insure that the scenario selected by the officer fit the requirements of a decision presenting a unique command challenge. Decision probes (Table 3) focused on specific components of the RPD and analytic models of decision making. Data from the application of these probes were used to construct an incident account, an example of which can be seen in Appendix B for the urban fireground knowledge base. The incident account was considered the primary data source for further coding and analysis of the scenario.

Table 3

Decision Probes for Wildland Fireground Commander Knowledge Base

Probe Type	Probe Content
Goal elicitation	What was the command objective?
Information sources	What information relevant to the decision was available and how was it obtained?
Information confidence	How confident was the officer in the information?
Decision options	What other options were considered?
Decision basis	How was this option selected? What rule was being followed?
Conscious assessment	How conscious was the officer that he was making a decision?
Critical experience	What specific training or experience was necessary to make this decision?
Novice mistakes	How might a less experienced officer have behaved differently? Where are mistakes most likely?
Potential aiding	If the decision was not the best, what training, knowledge, or information was missing which could have aided the decision?
Time pressure	How much time pressure was involved in making this decision?
Decision time	How long was taken in actually making the decision?

CDM Products

The value of the CDM in applied settings is the nature of the products it generates. In order to be useful it must be able to present an accurate description of the bases of expert decision making. Not only must it be able to elicit knowledge, but it must be able to represent the data and communicate the data to personnel who want to apply the products to organizational goals.

We are still learning how to develop and present CDM data. The following discussion will describe types of products that we have generated with the CDM.

Critical Cue Inventory. The Critical Cue Inventory (CCI) is an inventory of the information and perceptual cues that the expert used in assessing a situation. Specific probes for critical cues are shown in Table 3 under information sources. Obtaining critical cues for decisions based on the RPD model was particularly difficult just as it is difficult for virtually every knowledge elicitation method. The recognitional nature of these decisions made information cues less accessible to consciousness. Where time pressures were less extreme and opportunities for deliberation were available, critical cues could be obtained more easily.

The strongest use of a CCI thus far has been in our study of paramedics. The CCI consisted of the cues in context, for recognizing heart attack victims during and prior to their showing standard symptoms. This CCI is serving as the basis for a medical videotape to train perceptual discriminations.

Situation Assessment Record. Because the RPD model treats decision making as a form of complex pattern matching, much of the expertise elicited appears as situational assessment. This reflects an expert's understanding of the dynamics of a particular case and is the basis for the ability to recognize cases as examples of standard prototypes. Often, an initial situation assessment is maintained throughout an incident, with new information serving to elaborate on what was originally known. Changes in situational assessment are defined as the receipt of new factual information, or a perceived change in the nature of observed cues.

The documentation of situation assessment was done by a checklist of dimensions reflecting different classes of causal factors that are learned and interpreted by the expert to suggest and constrain courses of action. The dynamics of situation assessment were inferred from the expertise needed to interpret new facts and perceive changes. The elaboration of the required expertise took the form of a knowledge analysis based on critical cues (see Critical Cue Inventory).

An example of a checklist used to develop a Situation Assessment Record for the fireground commander knowledge base is shown in Table 4.

The nine situation assessment dimensions in Table 4 reflect the features used by fireground officers in their incident accounts and were developed by referring to the Critical Cue Inventory.

Table 4

Situation Assessment Checklist for Urban Fireground Commander Knowledge Base

Dimension	Cues
Problem	a. smoke: color, amount, toxicity b. fire: amount, location c. explosior potential d. chemicals e. rate of change
Structure	a. type: factory, house, office, vehicle, etc. b. materials: wood, brick c. architecture: special features d. age
Problem x Structure	a. seat of fire b. possibilities for movement
Weather	a. temperature b. moisture c. wind: velocity, direction
Risk to Life	a. direct cues b. knowledge of potential risk c. special populations: elderly, disabled etc.
Risk to Firefighters	
Nature of Attack	a. progress b. hindrances
Resources	a. what is available b. what is needed c. special needs
Goals Assessment	a. search and rescue b. fire control c. property conservation

The Situation Assessment Record (SAR) is a procedure for describing the dynamic shifts during an incident. At each point, the critical cues are defined along with current goals. The Decision Points (DPs) probed are represented at that level. As new information is received, the situational awareness changes: in a SA-Elaboration the goals become more explicit, and in a SA-shift the goals are rejected and replaced.

Table 5 shows a schematic for a SAR. In Table 5, SA-1 is the situation awareness when the incident began and includes the cues and goals upon which decision points (DPs) 1 and 2 were based. The receipt of additional information led to a deeper situation awareness (SA-2) which permitted elaboration of the original goals. Later information, not in keeping with the initial cues, caused an alteration in the situation awareness (SA-3) leading to the rejection of the original goals and the establishment of new ones. The final three decisions were made based upon this altered situation awareness.

We have used the SAR primarily for basic research purposes, to discuss various cases and dynamics.

Table 5

Situational Awareness Record

SA-1 Cues/Goals			
	DP - 1 Options Rationale	DP - 2 Options Rationale	
SA-2 Elaboration - Cues/Goals			
	DP - 3 Options Rationale		
SA-3 Shift - Cues/Goals			
	DP - 4 Options	DP - 5 Options	DP - 6 Options
Rationale Rationale Rationale			

Another method of representing SA is a more formal symbolic representation of SA aspects. Figure 5, from a study by Brezovic, Klein, & Thordsen, 1987, on tank platoon leaders illustrates this method. It shows typical cycles of situation aspect category use that both student tank platoon leaders and instructors went through at particular decision points. The DPs are shown in the diamond-shaped symbols. This has been developed as an effective means of letting readers visualize the difference in SA between experts and novices.

A last method is simply to present the SA of experts and novices side by side, in verbal format, for the same decisions. This was also done for the study of tank platoon leaders, and was very effective in portraying the increased sensitivity of experts to contextual nuances.

<u>Decision Point Analysis</u>. The purpose of the Decision Point Analysis was to create a detailed account of the decisions or course of action that could have been taken by the expert. Each decision point in the Incident Account was

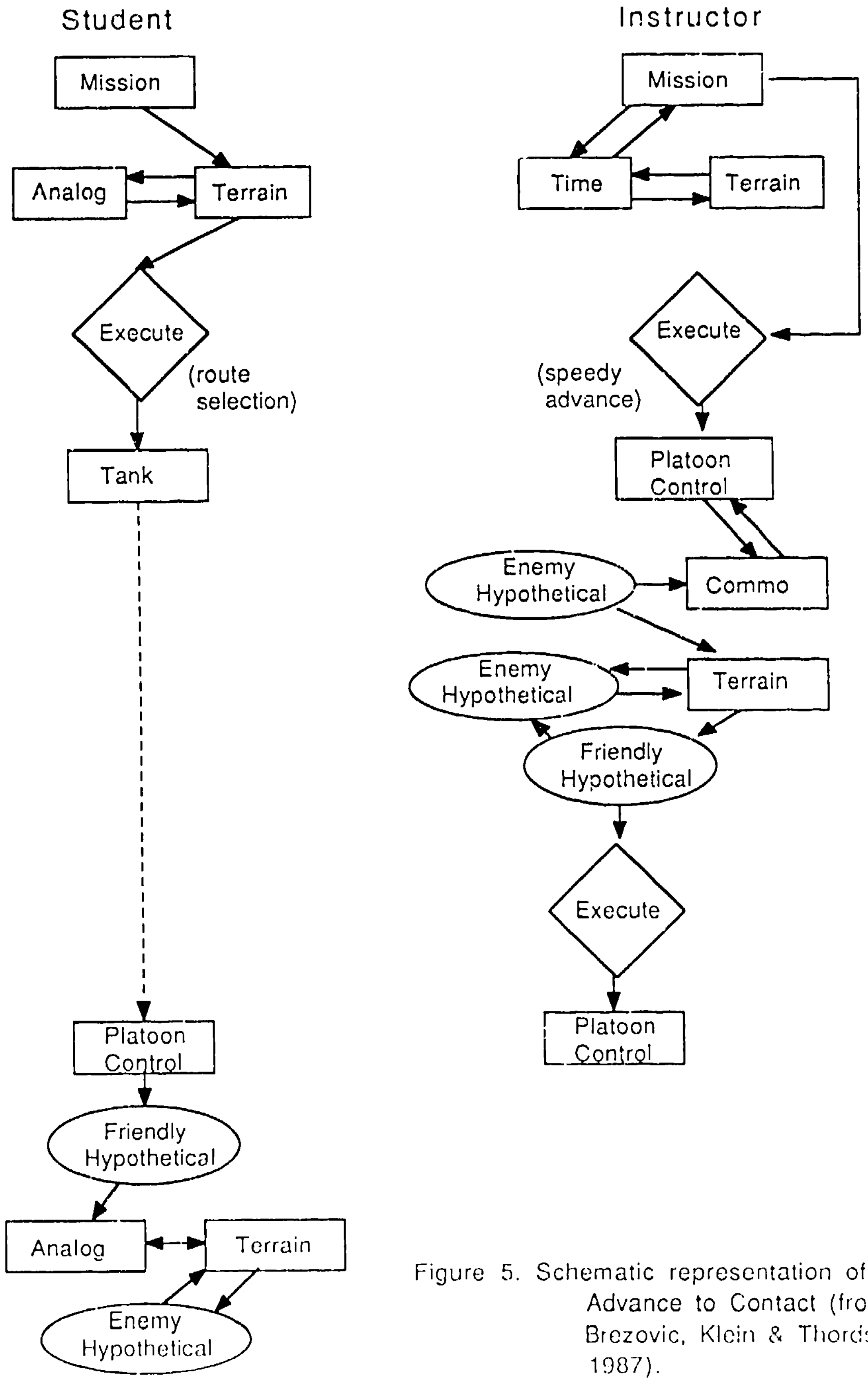

Figure 5. Schematic representation of Advance to Contact (from Brezovic, Klein & Thordsen, 1987).

characterized in terms of a number of dimensions and according to a set of specific criteria for each dimension. The analysis considered four distinct components of decision making: evaluation of options, implementation of actions, process characteristics, and timing. These are described below. It must be remembered that the Decision Point Analysis seems to be primarily useful for basic research into decision making.

1. Evaluation of options. The evaluation of options component examined in detail the manner in which options or possible courses of actions were generated and evaluated. Evaluation was regarded as a cognitive feasibility testing (see Recognition-Primed Decision Model), either by comparison with other options or actions in a queue, or according to a set of evaluation criteria. A distinction was made between actions implemented as a direct result of situation assessment (i.e., recognitional decisions) and actions implemented after a conscious deliberation that compared multiple options (i.e., analytic decisions).

2. Implementation of actions. This component detailed the expert's deliberations concerning both the method and timing associated with taking a chosen action. Judgmental processes dealing with strategic and tactical aspects of action implementation were documented, along with critical situation cues used to determine the appropriate timing of an action. These qualities of judgment are extremely critical in the performance of dynamic tasks and are typically learned only through extensive experience. Analyses of the components of these judgments can be used to develop refined training models by pinpointing the cues experts utilize. These can be taught in the context of simulation exercises or as part of narratives illustrating particular instances of decision expertise.

3. Process characteristics. This portion of the analysis was designed to characterize the nature of the mental processes used by experts in managing the decision problem. Delineations were made between recognitional, unconscious processes, and analytic, highly structured processes and other heuristics processes such as analogical reasoning and mental imagery.

4. Timing. This component of the analysis documented the amount of time taken by the expert to perform assessment and evaluation functions. It is used as a means of gauging the opportunity for aiding task components. Ongoing tasks (e.g., monitoring) or tasks for which significantly large amounts of time are afforded are more amenable to the application of aids than are tasks that require quick responses in a short time frame.

<u>Examples of Decision Point Analysis</u>. The decision point analysis structure outlined above was applied to the study of both the urban and wildland fireground knowledge domains. Tables 6 and 7 illustrate how the decision point analysis was implemented in analyzing the incident accounts for each of these two tasks. Complete details of the data collection and substantive findings appear in Klein, Calderwood, and Clinton-Cirocco (1985) for the urban fireground command example and in Klein and Taynor (1987) for the wildfire command example. In both cases, the analysis was broken down into separate stages, each involving a set of categories and criteria. The stages are instantiations of the Decision Point Analysis components outlined above; the categories and criteria reflect components of the RPD and analytic decision models detailed earlier.

Table 6

Model of Decision Point Analysis for Urban Fireground Commander Knowledge Base

Stage	Categories	Criteria
COMMAND LEVEL	Command Level	Decisions affecting the overall attack plan.
	Non-command level	Monitoring and coordinating activities.
SITUATION ASSESSMENT	Serial Situation Evaluation	Deliberation about nature of problem; tests possible hypotheses serially.
	Concurrent Situation Evaluation	Deliberation about nature of problem; simultaneously considers all aspects.
OPTION EVALUATION	Serial Option Evaluation	Deliberation about options; evaluates options serially.
	Concurrent Option Evaluation	Deliberation about options; simultaneously considers possible options.
OPTIONS CONSIDERED	Standard	Options explicitly taught or written down
	Typical	Contextually bound; apparent to other similarly competent officers.
	Constructed	No standard operating procedures could be described; creative problem solving required.
OPTION SELECTION	Abstract process	Decision making strategy could be explicitly taught; fairly easily articulated.
	Analogue process	Highly experimential process; highly context sensitive.
IMAGERY	Used/Not used	Evidence of imagery related processes; mental pictures.
DECISION TEMPO	Time for situation assessment, action evaluation, implementation evaluation	Self reports of clock time.
STRESS LEVEL	High/Low	Self reports of stress.
ATTACK MODE	Offensive/Defensive	Self reports of strategic orientation.

Table 7

Model of Decision Point Analysis for Wildfire Commander Knowledge Base

Stage	Categories	Criteria
DELIBERATION ABOUT THE SITUATION	Automatic Situation Recognition	Immediate recognition of situation
	Verify Situation	Confirmation of situation using critical cues
	Monitor Situation	Monitors situation until an action needs to be taken; has identified situation and knows how it will develop.
	Serial Situation Evaluation	Deliberation about nature of problem; tests possible hypotheses serially.
	Concurrent Situation Evaluation	Deliberation about nature of problem; simultaneously considers all aspects.
DELIBERATION ABOUT POSSIBLE ACTIONS	Automatic Action Selection	Option or action is linked to the situation assessment.
	Verify Selected Action	Strong inclination to take an action but considers it before proceeding.
DELIBERATION ABOUT IMPLEMENTING ACTIONS	How Implement Action	Deliberation about how to implement action.
	When Implement Action	Deliberation about when to implement action.
OPTION EXECUTED	Typical	Contextually bound; apparent to other similarly competent officers.
	Constructed	No standard operating procedures could be described; creative problem solving required.
TYPE OF REASONING	Highly abstract	Decision making strategy could be explicitly taught; fairly easily articulated.
	Experience	Analog proacess; highly experiential; highly context sensitive.
	Blend of abstract and experience	
DECISION SPEED	Time-bound deliberation	Clock time
	On-going/no deliberation	Untimed
IMAGERY	Used/Not used	Evidence of imagery; mental pictures.
KEY INDICATORS	Critical cues	Key indicators that allowed the expert to make the decision.

In addition to the analytic components central to the Decision Point Analysis approach, other components were added according to attributes specific to each task. For example, the urban fireground command task was much faster paced than the wildland fireground command task. Consequently, the manner of attack took on special significance to the command expert and required attention in the analysis.

Case Studies. One of the most vivid outputs of the CDM is the stories themselves, the brief anecdotes of how critical decisions were understand and handled. We have found that many people engaged in administrative tasks much prefer these concrete examples and case studies to abstract tabulations of decision types, etc. These data should not be overlooked as training materials. We have used brief case studies to describe the expertise of data analysts to a sponsor investigating the possibility of building an expert system. As a result of the sponsor's new understanding, the expert system project was abandoned.

Summary

There are many ways to implement the CDM. In each of them, the focus is on critical decisions and probes of situational assessment and decision points. But the variety of applications can differ greatly. We have used the CDM to explore the flow of decision making within single events. We have used it to examine specific decision points, comparing experts and novices. We have probed experts alone at specific decision points without gathering situational assessment record data or looking at the dynamic flow of the incident.

We have found that the method generates good reactions from subjects and users both. The subjects seem to prefer it to other psychological testing methods that rely on paper and pencil tasks or probe for more and more rules. The CDM seems to do a better job of respecting their expertise. It lets them tell their stories, especially the ones they are most proud of or interested in. Often we encounter suspicion when we enter a new domain but by the time we have done some probing their reactions transform into curiosity, and then pride. People we were not planning to interview have come up and volunteered to participate, because they want to have their expertise recorded. Virtually every time we have finished a report and shown it to the subjects, they have been pleased with the accuracy of the account, and on some occasions have shown it to other people whom they want to understand better the way they think.

Compared to the more controlled protocol analysis methods, the CDM has advantages by probing actual incidents, within a context, and by trying to reflect the most important aspects of that context.

DECISION-AIDING AND TRAINING IMPLICATIONS

Identifying Decision-Aiding Opportunities

Decision-aiding technologies are those that provide assistance and support to decision making either by improving (a) the knowledge and information bases used as inputs to a decision-making effort, or (b) the processes used as part of the structuring, combining, and analysis of that information.

The design of decision aid and support technologies generally begins with a model of the task to be performed. Often the model selected is a normative one, such as a decision analytic model. The rationale generally given for this choice is that such models provide optimal solutions to a problem and cannot be outperformed by humans' natural abilities. An alternative and arguably more accurate reason is that systems designers are often far removed from the substantive aspects of the tasks with which they deal. Tasks are modeled in the abstract according to highly generalized principles that are familiar to them and that obey known rules and relationships. Problems modeled in this way are often more elegant and tractable intellectually than is the case if the problem is dealt with in terms of its actual substance. Unfortunately, however, the task is then fit to the model rather than the model to the task. Consequently, opportunities for aiding and support are sought by looking to the designer's abstract model rather than to the task's substantive elements.

The knowledge elicitation approach presented here is intended as a first step toward remedying this deficiency in the system design process by developing a set of methodologies for identifying critical elements in task performance in terms that directly relate to the task as defined by those who perform it. This approach argues that experts in a task possess an understanding of the critical elements in task performance which, if properly modeled, can provide both structural and substantive specifications for the task. Those specifications can be used as a basis for selecting task components that are potentially amenable to aiding. These could be selected by applying a number of criteria, including:

1. criticality: The more critical a component is for overall task performance, the more important it becomes to have the component performed accurately and reliably. Identification of criticality can be done by expert judgment or by modeling the task in terms that emphasizes the linkage between task components. Components that are important precursors or inputs to other components are also more critical.

2. temporal sufficiency: A great number of decision tasks must be performed within a dynamic context. Consequently, constraints exist on the amount of time available for the performance of a task component before a window of opportunity has passed. In such contexts, decision performance can be improved if the decision maker is afforded a greater degree of temporal flexibility by the application of a decision aid. For example, a critical component in many decision domains is situation assessment; the sooner a situation can be accurately assessed, the sooner appropriate actions can be taken. In fast-paced

environments such as fire control, a relatively small improvement in the amount of time required to perform situation assessment can have a large impact on the quality of a decision maker's performance.

3. quality of human performance: Even experts are not always exceptional at all aspects of the tasks they perform. Aspects of the environment in which they work can force upon them tasks that are inherently difficult for humans to do and make them prone to errors. Highly trained, expert pilots, for example, are required to execute pre-flight checklists to insure that important task elements will not be overlooked due to limitations in their ability to remember the large number of details associated with preparing an aircraft for flight.

4. technical achievability: Not all identified opportunities for decision aiding will be technically feasible. The technical achievability of decision aids varies along a continuum. Highly generalized process aids are the most difficult to construct. This category includes aids intended to enhance cognitive processes such as generation of decision options. In general, the more domain-independent a decision aid becomes, the more difficult it is to technically execute. Highly achievable aids, in a relative sense, are those that present, collate or summarize information relevant to specific aspects of task performance. These include aids designed to enhance perceptual organization (e.g., displays), provide databases of technical information, or assist in preplanning for decision making.

These criteria lead to the following kinds of questions which if included as part of the knowledge elicitation approach described in this report can be useful in identifying specific decision aids and support elements.

Task difficulty:

Given a model of the task, what task elements does the expert perform well? What elements are performed poorly?

What elements of the task does the expert say are difficult to perform?

What are the specific sources of stress in task performance? Can the stress be relieved by support and training?

What would the expert like in the way of assistance and support?

Time requirements:

Where is time a critical factor?

How can the tempo of the task be reduced?

Information requirements:

What data bases are required?

Are they accessible in a meaningful time frame?
Where is information highly time sensitive?

Where is information insufficient?

Where is there more information than can be adequately synthesized by the expert?

Where can information be made more meaningful? Translated into knowledge?

Communication requirements:

What messages need to be sent to others?

In what time frame?

How difficult is this?

What languages or codes are used?

How can they be improved?

Tool use:

What tools are currently in use?

How are they used?

How can their use be improved? Alteration? Instructions? Training?

Are there tools that can be adapted to other aspects of the task?

Are there tools from other tasks that can be adapted to the task?

Identifying Opportunities for Improvement of Training

An important question to ask of any training program is why it is being conducted. A proper formulation of the basic goal of training is important in identifying opportunities for how it might be improved. One important goal of training is preparation for further training. This goal is often emphasized in programs that prepare individuals for on-the-job training in the field under the supervision of those who have achieved a level of expertise. For example, an automobile mechanic might undergo a training program that emphasizes familiarity with automobile nomenclature as preparatory to learning troubleshooting procedures.

A second goal of training is proficiency in a skill prior to performance. Sometimes people are required to perform tasks only very rarely, or perhaps only once. Emergency procedures for aircraft pilots are one example, high-risk combat missions are another. In these kinds of situations, there is no opportunity for remediation or for close supervision by experts. The individual must learn the task to a high level of proficiency outside the real world context. Training for this mode of performance is often done in highly realistic exercises (e.g., flight simulation) with expensive training aids. Therefore, this kind of training places a strong emphasis on the creation of realistic training environments. Enhancement of environmental fidelity constitutes a major category of improvements to a training program of this kind.

A third goal of training is sustainment. In many tasks, some skills are applied more frequently than others. People often lose proficiency for skills that are applied infrequently. Remediation is a means of improving proficiency in skills that have lost their sharpness over time. Improving this category of training can be done by developing better techniques for identifying reductions in skills proficiency. Non-intrusive measures of skills proficiency that can be applied in the field are one possibility. Observational or interview techniques that identify, post-hoc, the types of errors made in task performance are another.

Traditional approaches to building training programs usually begin with a model of instructional system design (ISD). Though many variants on ISD exist, all generally share some common elements. The design process begins with an analysis stage in which training requirements are defined (e.g., needs, constraints, target population), an instructional program is set out (e.g., goal definition, task specification, media requirements, performance levels), and existing resources for training are surveyed and evaluated for suitability. In the second stage, the training program is designed; this includes definition and specification of tasks, instructional objectives, entry knowledge, learning-task sequencing, assessment standards, and instructional materials. In the final stage, the training program is developed: learning activities are prepared, test items are developed, and instructional goals are reviewed and modified.

This approach is essentially a bottom-up orientation to instruction. It assumes that mastery of a skill must be achieved by incremental attainment of lower skill levels. Expertise is defined as an aggregation of knowledge through successful accomplishment of subordinate skills. It tacitly assumes that skilled performance is decomposable into ever finer increments of knowledge that can be trained independently. To its advantage, the ISD approach can be employed successfully to build training programs for basic skills, and deliver training cost-effectively to large numbers of individuals. However, at upper skill levels, it is much less attractive as a design approach (Klein, 1978). Expert skills are difficult to decompose according to an ISD format, performance requirements for testing and evaluation are extremely difficult to define, and learning objectives are difficult to specify. The phenomenological nature of expert performance requires that a different tack be taken to bridging the gap between the well-schooled novice and the well-experienced expert.

An alternative approach to ISD begins with the assumption that expertise is experiential in nature. Over a long period of time, the expert has accumulated a vast storehouse of knowledge concerning situations that might be faced, actions that are effective, and judgmental and decision-making strategies that work. Some of this information can be articulated easily and directly, and some of it cannot. The novice, rather than being schooled in the elements of the task, interacts directly with the expert in an experiential context, observes what the expert does, questions the expert when appropriate, receives the expert's rebukes when mistakes are made, and comes to learn the task holistically from performance and experience. This is not a new approach; it embodies elements of what has at other times been called the "socratic" teaching method. The novice learns from the master in a one-on-one (or one-on-few) format. The "Suzuki" method of musical instruction is somewhat akin to this approach; music theory instruction is shunned in favor of immediate, student-motivated performance, and the "expert" is present to guide the development of promising skills.

Building an instructional program on these principles requires an understanding of the knowledge that sets an expert apart from others. It proposes that one avenue to achieving expertise is to teach it directly. By careful study of expert solutions to a wide range of situations, a "data base" of expertise can be developed for a task that is a useful resource for transferring upper-level knowledge to others. That data base should include not only the actions that the expert took in a given situation, but also the rationale behind the actions and the mental deliberations involved in assessment and choice. The knowledge elicitation methodology described here has as one of its goals the development of such a knowledge base. It is premised on the alternative instructional assumptions outlined above and is designed to tap distinctive elements of expertise that define global qualities such as leadership, executive, and command abilities. These qualities are essentially unteachable by the traditional ISD approach.

SUMMARY

This report has presented a general approach for conducting knowledge elicitation of experts in the context of command decision making. To date, the majority of knowledge base development processes have emphasized the use of highly structured methodological tools designed to fit conveniently into the knowledge representation schemes of expert systems and instructional design technologies. To their credit, they lead to knowledge bases that are readily operated on by computer technologies, both hardware and software. To their detriment, they fail to preserve the contextual aspects of decision making and call upon experts to relate their experiences in unnatural and often meaningless ways.

The approach developed here was derived from cognitively based models of decision making rather than from the needs of a knowledge representation scheme. Grounded in the principles of process tracing and protocol analysis, the methodology elicits specific components and focuses on critical components of information and information utilization that define expertise.

REFERENCES

Boose, J. H. (1985). A knowledge acquisition program for expert systems based on personal construct psychology. Int. J. Man-Machine Studies, 23, 495-525.

Brezovic, C. P., Klein, G. A., Thordsen, M. (1987). Decision making in armored platoon command (MDA903-85-C-0327). Alexandria, VA: U.S. Army Research Institute.

Buchanan, B. G., & Shortliffe, E. H. (1984). Rule-based expert systems. Reading, MA: Addison-Wesley Publishing Co.

Calderwood, R., Crandall, B., & Klein, G. A. (1987). Expert and novice fire ground command decisions (MDA903-85-C-0327). Alexandria, VA: U.S. Army Research Institute.

Crandall, B., & Klein, G. A. (In preparation). Critical cues for MI and cardiagenic shock symptoms. Working paper. Yellow Springs, OH: Klein Associates Inc.

Crandall, B., & Klein, G. A. (1988). Key components of MIS performance. Report to the Standard Register Company. Yellow Springs, OH: Klein Associates Inc.

de Groot, A. D. (1978). Thought and choice in chess (2nd ed). New York: Mouton.

Dreyfus, H. L. (1979). What computers can't do: The limits of artificial intelligence (2nd ed). New York: Harper & Row.

Dudycha, L. W., & Naylor, J. C. (1966). Characteristics of the human inference process in complex choice behavior situations. Organizational Behavior and Human Performance, 1, 110-128.

Easterby-Smith, M. (1980). The design, analysis and interpretation of repertory grids. International Journal of Man-Machine Studies, 13, 3-24.

Ericsson, K. A., & Simon, H. A. (1980). Verbal reports as data. Psychological Review, 87, 215-251.

Ericsson, K. A., & Simon, H. A. (1984). Protocol analysis: Verbal reports as data. London: The MIT Press.

Fischhoff, B., Goitein, B., & Shapira, Z. (1982). The experienced utility of expected utility approaches. In N. Feather (Ed.), Expectations and actions: Expectancy-value models in psychology (pp. 315-339). Hillsdale, NJ: Erlbaum.

Fischhoff, B., & MacGregor, D. (1982). Subjective confidence in forecasts. Journal of Forecasting, 1, 155-172.

Fischhoff, B., Slovic, P., & Lichtenstein, S. (1978). Fault trees: Sensitivity of estimated failure probabilities to problem representation. Journal of Experimental Psychology: Human Perception and Performance, 4, 330-344.

Fischhoff, B., Slovic, P., & Lichtenstein, S. (1980). Knowing what you want: Measuring labile values. In T. Wallsten (Ed.), Cognitive processes in choice and decision behavior (pp. 117-141). Hillsdale, NJ: Erlbaum.

Flanagan, J. C. (1954). The critical incident technique. Psychological Bulletin, 51, 327-358.

Geiselman, R. E., Fisher, R. P., MacKinnon, D. P., & Holland, H. L. (1985). Eyewitness memory enhancement in the police interview: Cognitive retrieval mnemonics versus hypnosis. Journal of Applied Psychology, 70, 401-412.

Gettys, C., & Fisher, S. D. (1979). Hypothesis plausibility and hypothesis generation. Organizational Behavior and Human Performance, 24, 93-110.

Hammond, K. R., Hamm, R. M., Grassia, J., & Pearson, T. (1984). The relative efficacy of intuitive and analytical cognition. Boulder, CO: Center for Research on Judgment and Policy (see pg. 20).

Hayes, J. R. (1982). Issues in protocol analysis. In G. Ungson and D. Braunstein (Eds.), Decision making: An interdisciplinary inquiry (pp. 61-77). Boston, MA: Kent Publishing Co.

Hayes-Roth, F., Waterman, D., & Lenat, D. B. (1983). Building expert systems. Reading, Mass.: Addison-Wesley Publishing Co.

Howell, G. E. (1984). Task influence in the analytic intuitive approach to decision making final report (ONR Contract N00014-82-C-001 Work Unit NR197-074). Houston, TX: Rice University.

Hursch, C. J., Hammond, K. R., & Hursch, J. L. (1964). Some methodological considerations in multiple-cue probability studies. Psychological Review, 71, 42-60.

Kahneman, D., & Tversky, A. (1972). Subjective probability: A judgment of representativeness. Cognitive Psychology, 3, 430-454.

Keeney, R. L., & Raiffa, H. (1976). Decisions with multiple objectives: Preferences and value tradeoffs. New York: Wiley.

Klein, G. A. (1978). Phenomenological vs. behavioral objectives for training skilled performance. Journal of Phenomenological Psychology, Fall, 139-156.

Klein, G. A., & Brezovic, C. P. (1986). Decision engineers and the design process: Decision strategies and human factors literature. Proceedings of the 30th Annual Human Factors Society, 2, 771-775. Dayton, OH: Human Society.

Klein, G. A., & Calderwood, R. (1986). Recognitional and calculational decisions in skilled chess performance. Yellow Springs, OH: Klein Associates Inc.

Klein, G. A., Calderwood, R., & Clinton-Cirocco, A. (1985). Rapid decision making on the fire ground. Yellow Springs, OH: Klein Associates Inc.

Klein, G. A., Calderwood, R., & Clinton-Cirocco, A. (1986). Rapid decision making on the fire ground. Proceedings of the 30th Annual Human Factors Society, 2, 576-580.

Klein, G. A., Maher, Zakay, D., & Kessel, (1987). Knowledge engineering analysis of data extraction strategies (Technical Report). Yellow Springs, OH: Klein Associates Inc.

Klein, G. A., & Peio, K. J. (1982). The use of a prediction paradigm to study proficient performance. Yellow Springs, OH: Klein Associates Inc.

Klein, G. A., & Taynor, J. (1987). Modeling the expertise of wildland fireground commanders. Yellow Springs, OH: Klein Associates, Inc.

Klein, H. A. & Klein, G. A. (1978). Perceptual/cognitive analysis of proficient cardio-pulmonary resuscitation (CPR) performance. Yellow Springs, OH: Klein Associates Inc.

Lichtenstein, S., Fischhoff, B., & Phillips, L. (1982). Calibration of probabilities: State of the art to 1980. In D. Kahneman, P. Slovic & A. Tversky (Eds.), Judgment under uncertainty: Heuristics and biases (pp. 306-334). New York: Cambridge University Press.

Lichtenstein, S., & Slovic, P. (1971). Reversals of preference between bids and choices in gambling decisions. Journal of Experimental Psychology, 89, 46-55.

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

Pitz, G. F., Sachs, N. J., & Heerboth, J. (1980). Procedures for eliciting choices in the analysis of individual decisions. Organizational Behavior & Human Performance, 26, 396-408.

Polany, M. (1966). The tacit dimension. Garden City, NY: Doubleday & Co., Inc.

Raiffa, H. (1968). Decision analysis. Reading, MA: Addison-Wesley Publishing Co.

Rosch, E., & Mervis, C. B. (1973). Family resemblances: Studies in the internal structure of categories. Journal of Cognitive Psychology, 7, 573-605.

Rouse, W. B. (1978). Human problem solving performance in fault diagnosis task. IEEE Transactions on Systems, Man, and Cybernetics, SMC-8, No. 4, 258-271.

Rouse, W. B., & Morris, N. M. (1986). On looking into the black box: Prospects and limits in the search for mental models. Psychological Bulletin, 100, 349-363.

Schiffman, S. S., Reynolds, M. L., & Young, F. W. (1981). Introduction to multidimensional scaling. New York: Academic Press.

Schvaneveldt, R. W., Durso, F. T., Goldsmith, T. E., Breen, T. J., Cook, N. M., Tucker, R. G., & De Maio, J. C. (1986). Measuring the structure of of expertise. International Journal of Man-Machine Studies, 23, 699-728.

Shaw, M. L. G., & Thomas, L. F. (1978). FOCUS on education -- an interactive computer system for the development and analysis of repertory grids. International Journal of Man-Machine Studies, 10, 139-173.

Simon, H. A. (1955). A behavioral model of rational choice. Quarterly Journal of Economics, 69, 99-118.

Svenson, O. (1979). Process descriptions of decision making. Organizational Behavior and Human Performance, 23, 86-112.

Svenson, O. (1986). On the collection and analysis of verbal protocols in process studies of judgment and decision making. In H. Montgomers & O. Svenson (Eds.), Process and structure in human problem solving.

Taynor, J., Klein, G. A., & Thordsen, M. (1987). Distributed decision making in wildland firefighting. Yellow Springs, OH: Klein Associates Inc.

Thordsen, M., Brezovic, C. P., & Klein, G. A. (In preparation). Knowledge elicitation study of command-and-control decision making. Working Paper. Yellow Springs, OH: Klein Associates Inc.

Tversky, A., & Kahneman, D. (1971). Belief in the law of small numbers. Psychological Bulletin, 2, 105-110.

Tversky, A., & Kahneman, D. (1973). Availability: A heuristic for judging frequency and probability. Cognitive Psychology, 4, 207-232.

Tversky, A., & Kahneman, D. (1974). Judgment under uncertainty: Heuristics and biases. Science, 185, 1124-1131.

Tulving, E. (1972). Episodic and semantic memory. In E. Tulving & W. Donaldson (Eds.), Organization of memory. New York: Academic Press.

von Clausewitz, C. (1943). On war. Princeton, N. J.: Princeton University Press.

Waterman, D. A. (1986). A guide to expert systems. Reading, Mass.: Addison-Wesley Publishing Co.

Winston, P. H. (1984). Art*i*ficial intelligence. Reading, MA: Addison-Wesley Publishing Co.

Zakay, D., & Wooler, S. (1984). Time pressure, training, and decision Ergonomics, 2, 273-284.

APPENDIX A

INTERVIEW GUIDE FOR THE URBAN FIREGROUND KNOWLEDGE ELICITATION

FIREGROUND DECISION INTERVIEW GUIDE AND DATA RECORD

Department: ______________________ Interview date: ____________

Conducted by: ____

Transcribed by: ____

Interview time/transcription time (in man hours) ________________

I. INTRODUCTION

Describe purpose of the study - learn about how experts such as command level fire officers make decisions under extreme time stress. Klein Associates (small industrial psychology company established in 1978). Interview will focus on decisions made at fires which were demanding from a command perspective. Approximately 2 hours to conduct interview.

II. BIOGRAPHICAL DATA

Name/rank:

Firefighting Experience (years, where, positions held, approx. dates):

Optional remarks (special training, job satisfaction, etc.):

III. INCIDENT DATA

A. Choice of Critical Incident

The incident may have been preselected as recent incident of interest. If not, officer should choose most recent incident which presented a challenge. In general, the more serious the incident, the more likely it will be that command, rather than procedure will play an important role. Any factors which make the incident exceptional in some way should be considered, such as risk to life, non-standard operations which were employed, mistakes which were made, etc. Errors in judgment may be particularly informative.

Note why incident was selected:

_____ Recent _____ High risk _____ Disappointment _____ Non-routine

_____ Other ____

B. Officer's Incident Account

Officer is asked to relate all the events from the time the alarm was received, focusing on his commands and critical decisions. This part of the interview should be unstructured to allow the perspective of the officer to emerge. Probes on decision making and timeline details should be carefully timed so as to interfere as little possible as with important points the officer wishes to make.

C. Probing for Specific Information

1. Timeline details. Because we are so interested in how time pressure affects decision making, we wish to gather as much information as possible as to the sequencing and duration of events occurring at the fireground. The timeline also functions to clarify and even aid the officer in recalling the incident. If this is difficult for the officer, stress that relative time information is more important than clock time. It may be possible to check certain details against supporting documents (some incidents are reported in timeline form as part of the incident report).

2. Decision/Command Probes

In general we are interested in finding out all that we can about how a fire ground commander makes decisions on the fire ground, including critical decisions which were faced, options considered at each decision point (why one was chosen, others rejected), strategies employed, critical information available and lacking, etc. During the interview, the interviewer will identify and probe key decisions which were made at the scene.

The Timeline can serve as a partial checklist for the type of information which we wish to have for each key command/decision. Additional probes are more or less relevant or important depending on the nature of the incident and the type of information revealed in the incident account. A number of probes for the type of information considered central to this study are listed on the next page. Although this list can serve as a checklist, it is best for the interviewer to be thoroughly familiar with these probes and to judge when they are most appropriate.

Keylabel	KEY COMMAND/DECISIONS
GOAL	What was the decision/command objective?
INFO/SOURCE	What information relevant to the decision was available and how was it obtained?
INFO/CONF	How confident was officer of information (low, med, high)?
OPTIONS	What other options were considered?
BASIS	How was this option selected? (What rule was being followed? If no rule, probe for use of analogy, scenario, etc.)
CONSCIOUS	How conscious was officer of making decision?
CRIT EXP	What specific training or experience was necessary to make this decision? (at what point in his career would the officer <u>not</u> have had the requisite knowledge to make a good decision).
MISTAKE	How might a less experienced officer have behaved differently? (Where are mistakes most likely?)
HELP	If decision was not the best, what training or knowledge or information was missing which could have aided the decision?
TIME PRESSURE	How much time pressure was involved in making this decision (scale 1-4)? 1 = very low, as low as ever experienced in an incident 4 = very high, as high as ever experienced in an incident
TIME	Estimate how long was taken in actually making the decision.

INCIDENT PROBES

ANALOGY	Pick the most similar/helpful case. Describe differences.
SCENARIO	Does the incident fit a standard scenario that you have ever seen or been trained for? (Probe basis for match, differences/modifications)
#INCIDENTS	How many incidents like this have you been involved in? overall _____ at a command level _____
STRATEGY	What were the overall strateg[illegible] employed (offensive/defensive during course of inc[illegible]nt?

CODING ABBREVIATIONS

Information

D = direct visual/auditory/smell
R = reconnaissance
P = preknowledge
Disp = dispatch
FF = other fire fighter
C = citizen report
Unk = if an event or piece of information was unknown at the time of its occurrence. On data record, note when the information became available

Strategy objectives

S/R = Search/rescue
FC = Fire control
PC = property conservation

Specific Goal

SZ = sizeup
M = manpower
E = equipment
Evacuation
Safety
Speed
Planning

Degree of Consciousness

Automatic
Some reflection
Conscious consideration of alternative

General Abbreviations

FF = fire fighter
FG = fire ground
FGC = fire ground commander
EMS = emergency medical service
EMT = emergency medical technician

APPENDIX B

EXEMPLAR INCIDENT ACCOUNT FROM THE URBAN FIREGROUND KNOWLEDGE BASE

INCIDENT ACCOUNT #1

TANKER TRUCK

This incident occurred 18 months ago. The incident involved an overturned tanker truck on fire. It had been carrying a full load of jet fuel, on an access ramp of an interstate highway during morning rush hour. The Chief had never been involved with a tanker fire before and this fire was particularly hazardous because of the existence of another tanker approximately fifty feet behind the overturned one.

Chief McW heard the dispatcher call sometime during the city's morning rush hours. Instantly recognizing the location as being within his area, he headed out in his car toward the area given by the original dispatcher call. The only information given by the dispatcher was that a tractor trailer was involved. On his way to the scene, the Chief saw a huge column of black smoke coming from the freeway at a location that was not the area of the original report. The location of the box alarm was some distance from the incident. The Chief acted on the visual cue of the smoke and arrived at an expressway off-ramp of a major interstate within two minutes of the call. Getting closer to the scene, the Chief saw a column of flame and citizens running from the scene, abandoning their cars. On his arrival he saw a tanker truck laying on its side, ruptured lengthwise, and engulfed in flame. "I immediately breathed a sigh of relief, because the danger of explosion was less than if it (the truck) was split in half," the Chief noted. A second tanker was about fifty feet from the overturned one. At the same time as this quick size-up, the Chief started toward a group of citizens who were helping the injured driver of the crashed vehicle away from the immediate area. While assisting in this evacuation, the Chief questioned the driver about his load and found that the tanker was carrying JP-4 jet fuel and had just been fully loaded. About thirty seconds had elapsed from the Chief's arrival.

The Chief then got on his radio and (1) corrected the address given by the original dispatcher call, (2) called for a rescue unit for the injured driver, (3) requested police action to stop the flow of traffic, and (4) called for a special firefighter unit that dispensed foam. By the time the call was completed, the first alarm units had arrived and were attempting to hook up to the nearest hydrant located some distance from the scene. A five-inch hose was going to be used from this source - a size that would drain the reservoir carried on an engine in about a minute. A smaller size hose was connected to the engine and was directed toward setting up a stream of water around the wrecked tanker; the hydrant supply was not available for about 15 minutes after the units had arrived. The Chief directed the streams to be set up for the protection of the firefighters and would not allow his men to advance on the fire until the protective streams were in place. He saw that the fire was well underway and pretty intense, but burning straight up and not threatening to expand much. The danger was that the saddle tanks of the tanker or other pockets of fuel would explode. A ladder truck was also directed to extend its ladder pipe and aim a water stream down on the scene. While this was being accomplished, the Chief sent the driver of the first fire truck on the scene down the ramp to check for occupants in the abandoned cars.

Two foam units then arrived, one at each end of the damaged tanker. The Chief coordinated their foam-dispensing operations so that the streams were at acute angles to each other. At this point the Chief felt that the situation was pretty much under control - but then a storm sewer behind him "blew," i.e., exploded into flame. He realized that burning fuel was now in the sewer system and recognized that this new aspect of the situation would exceed his span of control. He called for another alarm to be given. The next Chief arriving with these new units was tasked with removing the danger from the sewers while Chief McW left his attentions on the tanker.

The total time to containment was more than an hour. Chief McW was at the end of his shift during the mop-up phase of the tanker operations and decided to go home when all that remained was to right the truck.