Decision Models: Some Lessons From the Field

Gary A. Klein and Roberta Calderwood

***Abstract*—In recent years there have been a number of studies of decision making in naturalistic environments such as nuclear power plants, urban fire departments, command-and-control centers, and courtrooms. These studies provide a different perspective on decision strategies from laboratory studies that rely on naive subjects performing relatively context-free tasks under little time pressure and static task conditions. The field research paradigms generate a description of proficient decision making that emphasizes recognitional rather than analytical processes and challenges the "decision tree" representation of decision problems. That is, failure to use analytical methods is not a sign of incompetence since the recognitional strategies shown by decision makers seem reasonable and effective. Implications of these models for decision support systems and for training are discussed.**

I. Introduction

THE MAJOR ASSERTION of this paper is that the prevailing paradigms in decision research, based as they are in simplified and highly structured laboratory tasks, have limited utility in operational domains characterized by high time pressure, uncertainty, and ambiguity, continually changing conditions, ill-defined goals, and distributed decision responsibilities. Previous researchers have commented on the importance of making our research methods more responsive to applied needs [1]–[3]. In this paper we wish to reemphasize some of these issues from the point of view of decision research and practice. A review of studies based on naturalistic tasks supports the claim that with a different orientation to research tasks, a very different model of decision making emerges.

Specialists in decision research may argue that the field is now so broad that it is impossible to characterize any single prevailing paradigm. Our claim, however, is that beneath the apparent diversity, there is a common or root metaphor that has determined what variables are likely to be considered and even what questions are likely to be asked. This metaphor is concretely illustrated by the pervasiveness of the decision tree [4] as a "natural" format for representing and thinking about decisions in terms of the "moment of choice." Thus, it is natural to speak of a decision maker as "faced with alternatives," which can be specified as branches emanating from a single point in the tree. It is also natural to speak of the decision maker "considering the consequences" of each alternative in terms of an analysis of future states (odds/probabilities) weighed against alternative goals (preferences/utilities). Thus, the decision tree metaphor has tended to highlight the decision maker as analyst—generating alternatives, and evaluating outcomes. The decision event is the focus of concern. We believe it is this focus that has supported the view, most closely associated with Daniel Kahneman and Amos Tversky, of human decision makers as inherently biased and suboptimal [5]–[7]. It has been a logical step, therefore, to focus decision support on methods of debiasing judgments [7], [8], calibrating probability estimates [9], instructing in optimal combinatorial methods [10], and the like. The result has been myriad training programs, aids, and procedures that emphasize analytic approaches to decision making.

Manuscript received May 20, 1990, revised February 4, 1991. This work was supported by the U.S. Army Research Institute Contracts MDA903-85-C-0327, MDA903-86-C-0170, and MDA903-89-C-0032. This paper is an extended version of the earlier paper, which was awarded the Franklin V. Taylor Award for best paper, entitled "Decision models: Some lessons from the field," published in the 1987 Proceedings of the IEEE Systems, Man, and Cybernetics Conference, Washington, DC, 1987.

G. A. Klein is with Klein Associates Inc., Fairborn, OH 45324.

R. Calderwood is with Science Applications International Corporation, Albuquerque, NM 87106.

IEEE Log Number 9144661.

As people have tried to implement these programs and decision aids, it has become clear that analytic approaches are not very useful in the field. Application of inappropriate decision models can result in a variety of unfortunate outcomes, including training programs that are of limited benefit [11]. In the hands of an expert, decision analysis can be helpful in identifying factors entering into the decision, and in enabling decision makers to understand the differences in their goals and values. But expert decision analysts have also learned the boundary conditions and are careful not to push the methods beyond those boundaries.

Perhaps the most important mismatch between naturalistic decision making and classical decision models is the fact that the primary effort is usually not the moment of choice but rather in situation assessment. Situation assessment means identifying and clarifying the current state of the world including goals and assumptions, which Gettys [12] has called the "predecision processes." Decision models must include predecision processes in order to address the many pressing applied needs for real-world decision making. We may be more successful in applying training and decision support to these predecision aspects of decision making than to the moment of choice.

The objectives of classical decision theory were reasonable. If strategies such as decision analysis were general prescriptive models for making effective decisions, then training or aiding users to apply decision analysis would have resulted in an overall improvement in decision making. The improvements would have extended beyond any one specific context. Even though this gamble does not appear to have paid off, it was worth taking because the benefits far outweighed the risks. It was only by attempting such applications that it was possible to learn that the prescriptive strategies were not compatible with task demands in operational settings. The problem is that classical decision theory rests on some restrictive assumptions.

0018-9472/91$01.00 © 1991 IEEE

Classical decision theory has been particularly useful in presenting prescriptive guidelines for making better decisions (as opposed to descriptions of how people actually make decisions). Decision analysis and multiattribute utility analysis (MAUA), along with Bayesian statistics, are analytical tools that have been imported from mathematics, economics, game theory, and statistics for the normative goal of defining ideal decisions. For a given situation, once we understand what an ideal decision strategy is, we can pursue the prescriptive goal of improving the quality of decision making, as by helping people understand how to use base rates, or combine probability data, or carry out other analytical methods that should lead to better performance.

Classical decision models have a number of clear strengths. They are generic models that can be widely applied, so there is less need to be concerned with any one specific domain. The techniques ensure that all participants in a decision task will be speaking the same language, and using the same metrics. The techniques also lend themselves to being incorporated into decision aids as a way of framing and guiding the decision making. Therefore, in those situations where it is feasible to apply prescriptive models it is possible to produce powerful techniques for overcoming biases and improving decision quality. In short, when the assumptions underlying the models are met, then the techniques can ensure optimal selection between options.

Unfortunately, classical decision models depend on a number of problematic assumptions (e.g., [13]): that goals can be isolated, that utilities can be assessed independent of context, that probabilities can be accurately estimated, that choices, goals, and evidence are carefully defined, and that the utilities of an outcome are independent of other outcomes. Each of these assumptions seems difficult to meet in an operational environment. If the assumptions are rarely met, then these models may not be useful guidelines for making better decisions. Let us examine the assumptions.

1) Can goals be isolated? Classical decision theory assumes that goals can be stated clearly as an anchor for the analyses. For example, a MAUA might take as a starting point the goal of slowing down an enemy advance by denying the enemy the use of key roads. In actuality, this goal is linked to parallel and higher-order goals such as using the same resources in other ways, or planning for an eventual counterattack a few days later over those same roads. Obviously, it is risky to segment goals out of the larger context. On the other hand, it can be overwhelming for a battle manager to try to deal with the larger picture in analytical terms. In messy environments, goals are often interrelated in many different ways and it is dangerous to make simplifying assumptions in order to isolate goals to make the analysis work.
2) Can utilities be assessed independent of context? MAUA requires the decision maker to judge the value of different options using a general set of evaluation dimensions that can be applied to all the options. The problem is that even in a domain as restricted as chess, we do not see chess masters rating different moves using a standard set of dimensions (e.g., center control, potential for king-side attacks). Instead, the moves are evaluated in context, in terms of the specific advantages and disadvantages of the outcomes arising from each option. Expertise often enables a decision maker to sense all kinds of implications for carrying out a course of action within a specific context, and this sensitivity can be degraded by using generic and abstract evaluation dimensions.
3) Can probabilities be accurately estimated? Decision analyses require people to estimate the probability of occurrence for different branches of an option tree. We know how difficult it is to assign probability estimates even without time pressure and psychological stress. Experienced decision analysts recognize this as a limitation of the technique—operators will not be able to provide reliable data as inputs into the analyses. Furthermore, probabilities are more suited for decision tasks involving repeated occurrences of randomly generated events, rather than unique "one shot" events.
4) Will choices, goals, and evidence be clearly defined? Decision analysis models work best with bounded decisions such as which car to buy, or whether to proceed with a surgical operation. In contrast, for many decisions the end states are not well defined. For example, a fireground commander rushing to a fire would seem to have a simple, well-defined goal—put the fire out. But if the fire has spread too far, it may make more sense to just prevent it from spreading. The commander can call in for reinforcements, but has to be careful since that leaves the district vulnerable to other fires. The actual goal is to do the best job possible with the appropriate amount of resources, hardly a well-defined goal. In domains such as running a business or preparing for a battle, the goals are even less likely to be well-defined.
5) Are the utilities of outcomes independent of other outcomes? In order to analyze a decision, it is necessary to treat each of the outcomes separately, and to independently estimate the utilities of each outcome. This may be possible for static and limited tasks. But for complex operations the outcomes of different courses of action are interrelated so decision makers can't be trusted to provide useful outputs.

There are a number of serious disadvantages to misapplying prescriptive decision models. If the assumptions are not met, then the models cannot be trusted to provide useful inputs. And it seems clear that operational environments will rarely meet most of the assumptions listed above. Worse yet, by trying to force experienced decision makers to adjust to the needs of the prescriptive models we run the risk of degrading their ability to make use of their own experience. We can interfere with their proficiency. That is why it is important to understand the basis of decision expertise in order to enhance decision makers' abilities.

It may be that studies of decision making in naturalistic settings will be helpful in deriving decision models that are more representative. We will describe a set of studies we performed with firefighters and other command-and-control personnel, and then examine the findings of other researchers working in naturalistic settings. The intent of this paper is to identify alternative descriptions of decision making.

II. Recognitional Decisions

For the past several years, field experimentation has been underway to examine the decision making of urban fireground commanders (FGC's), using observations and retrospective accounts of actual emergency events [14]. Some examples of the types of decisions these commanders have to make include: whether to initiate search and rescue, whether to initiate an offensive attack or concentrate on defensive precautions, and where to allocate resources.

It was difficult to represent the phenomenological accounts of these decision processes in any meaningful way within the decision tree framework. Indeed, the fireground commanders resisted any attempt to characterize their roles in terms of "making choices," "considering alternatives," or "assessing probabilities." They instead saw themselves as acting and reacting on the basis of prior experience, planning, monitoring, and modifying plans to meet specific constraints. There was no evidence for anything approaching exhaustive option generation. Rarely were even two options concurrently evaluated, so that opportunities for tradeoffs between the utilities of outcomes were largely absent. There was no obvious way to apply the concept of optimal choice. Such a term was seen as potentially paralyzing action. Although the commanders could clearly recognize and admit to making mistakes, "workable," "timely," and "cost-effective" were much more meaningful criteria.

It was necessary to abandon the prevailing decision tree metaphor as overly restrictive and misleading. Within the decision tree metaphor, these firefighters were not making decisions. Yet the FGC's were clearly encountering choice points during the course of an incident. That is, there was an awareness that alternative courses of action were possible. However the FGC's insisted that they did not deliberate about the advantages and disadvantages of the different options.[1] Instead, they relied on their ability to recognize and appropriately classify a situation to generate a typical way of reacting. If time pressure was extreme, they might just carry out this action. If possible, they would evaluate the option by examining its feasibility. Often imagery would be used, to "watch" the option being implemented, seeing if anything went wrong. If problems were found, then the option might be modified, or rejected altogether, and a next most typical reaction would be studied. Concurrent deliberation between options was present in a minority of the decision points that were analyzed [15].

[1] It is possible that the FGC's were contrasting alternatives (but at an unconscious level), or possibly the FGC's were unreliable in their reports. We have no way of demonstrating that the FGC's weren't contrasting alternative options, but the burden of proof is not on us. There is no way to prove that something isn't happening. The burden of proof is on those who wish to claim that somehow, at some level, option comparison was going on anyway. The reasons we believe that the FGC's were rarely contrasting options are: because it seems unlikely that people can apply analytical strategies in less than a minute (see [10]); because each FGC argued forcefully that s/he wasn't contrasting options; and because they described an alternative strategy that seemed to make more sense.

TABLE I
Frequency Distribution of Decision Categories Used ($n = 156$)

Type	Frequency
Concurrent Deliberation	29 (18.6%)
Recognition Match	127 (81.4%)
Analogue Match	3
Prototype Match	124

A total of 26 interviews were conducted, and 32 incidents were studied. For these incidents, 156 decisions were probed in sufficient detail to allow subsequent coding. Our approach was to elicit a retrospective protocol based on their moment-by-moment recollections of the events and their own thinking. Because of the problems of defining when a decision was made, we developed the notion of a "decision point" as being those instances when the FGC could, after the fact, agree that alternative courses of action had been possible, even if these alternatives were not considered at the time the action was initiated. The criterion for a decision was that a choice existed in the form of reasonable alternatives, regardless of whether the FGC considered any alternatives at the time the action was initiated. Table I presents the categories into which these decisions fell.

Less than 20% of the decisions involved concurrent deliberation, in which more than one course of action was considered and contrasted. The classical concept of a decision event, or moment of choice, is exactly this type of conscious evaluation of several different options, and yet it occurred infrequently in our sample. Moreover, our coding was designed to overcount such cases. If the interview showed any evidence that the FGC had compared two or more courses of action at any time during the incident, no matter how fleetingly, we counted this as a case of concurrent deliberation. Furthermore, our probing favored this category since we were expecting to find that the FGC's had narrowed the search down to no more than two or three alternatives, so the demand characteristic of the study was to elicit more instances of concurrent deliberation. Nevertheless, it was relatively infrequent in the cases we studied, and these were the most difficult and nonroutine incidents the FGC's could recall. One caveat is that we identified and probed a wide variety of decision points in each case, including very difficult ones as well as some more straightforward ones. If we had limited our analysis only to the most difficult and nonroutine decision points, the rate of recognitional decisions would have been less than 80%. Subsequent studies that did look only at the nonroutine decision points obtained rates of recognitional decision making between 50% and 70% (see [16] for a review). Our coding methods were shown to be highly reliable. Taynor *et al.* [17], found intercoder agreement to be between 87%–94%.

An example may be helpful. The commander of an emergency rescue team was called to the assistance of a woman who had fallen or jumped from a highway overpass. Instead of plunging to her death she had hit the metal struts supporting some signs, and was dangling on these when the team arrived. Two members of the crew climbed out and secured her arms and legs, but there was no way to gain any leverage and they needed to lift the woman to safety. The commander considered using a Kingsley harness that would snap onto her shoulders and legs, but, as he imagined carrying out such a rescue, he realized that it would be dangerous to attach the harness to a semiconscious woman lying on her stomach. She would have to be lifted to a sitting position, held still, and so on, while keeping everyone balanced on the narrow support poles. So he rejected this option. Next he imagined attaching the harness from the back, but as he played this out in his mind he saw the woman's back bending sharply as they lifted her up, and he rejected this option as well. He considered using a Howd strap instead of a Kingsley harness, but when he performed the mental simulation he found the same problems. Then he realized he could use a ladder belt for the rescue. The woman would just have to be lifted up an inch or two, the belt slid under her, the buckle attached, and a rope tied to a clasp on the belt to lift her to safety. The commander imagined this scenario a few times, and ordered his crew to use a ladder belt to make the rescue. Notice that several different options were considered but none was contrasted to another. Instead, each was examined for feasibility, and the first acceptable one was implemented.

These strategies have been described in a recognition-primed decision (RPD) model [16]. The model differs from standard models in that it begins with action based on the recognition of a situation as familiar or prototypical and posits a serial option evaluation strategy.

The RPD model is presented in Fig. 1. The model simply illustrates several types of recognitional decision strategies. A person understands a situation in terms of its familiarity to previous experiences. The judgment of familiarity to a given set of prior cases carries with it a recognition of goals that are feasible, cues that are relevant, expectancies to monitor, and actions that are plausible. The decision maker can use experience to generate a likely option as the first one considered. The evaluation of the option is through mental simulation to see if there are any pitfalls to carrying the option out. If these can be remedied, the option can be strengthened. If not, the option is rejected. If no pitfalls are envisioned, the option can be used.

This model has a number of distinctive features. It asserts that experienced decision makers can directly generate plausible options and do not have to randomly generate a large set of options to find one that is useful. It asserts that the decision maker can use a serial process of generating and evaluating options in turn, without having to generate a set of options to be contrasted with each other. The RPD model asserts that decision makers can use satisficing to select the reasonable option, rather than trying to optimize by picking the best possible option. It asserts that evaluation can involve mental simulation, thereby studying the option in the task context, rather than relying on MAUA to use abstract dimensions for contrasting different candidate options. It emphasizes the way the situation is understood more than the decision event of choosing between several alternatives. The RPD model presents a strategy that experienced decision makers can use to rapidly select an option, rather than having to complete an analysis in order to know how to act.

Within the context of this task environment, a recognitional strategy appears to make a great deal of sense. The proficient FGC's we studied could use their experience to generate a workable option as the first one they considered. If they had tried to generate a large set of options, and then tried to systematically evaluate these, the fires would have gotten out of control before they could make any decisions. Approximately 85% of the decisions probed were estimated to have been made in less than one minute. Evidence suggests that analytic decision strategies cannot be effectively accomplished in this time frame [10], [18], [19].

III. Field Research: Tasks and Paradigms

The RPD model was outlined to illustrate the limitations of the tree metaphor within the complexities of a real-world environment. In fact, it seems that the metaphor is consistently challenged whenever the focus of study has been on a specific task or application rather than on model testing. For example, several independent studies of command-and-control decision making have stressed the importance of recognition and situation assessment over the option evaluation strategies that are the standard focus of laboratory-based methods.

Limitations in the decision tree metaphor have been encountered by researchers studying command-and-control decision making, perhaps because it highlights the task features of time pressure, dynamic conditions, ill-defined goals, and high uncertainty. Wohl [20] presented a SHOR model (stimulus-hypothesis-option-response) that has been extremely influential in the development of subsequent field research and theorizing. Wohl showed the importance of tying theory to practice (his area of interest was Navy command-and-control decisions), and the SHOR model highlighted situation assessment to go along with option generation and evaluation. Lipshitz [21] has analyzed the decision protocols of Israeli command-and-control personnel and concludes that forcing the protocols into a tree structure results in numerous examples of what he terms a "pseudo double-option choice." This is the case where a choice point appears to have two alternatives, but the alternatives are neither logically nor psychologically equivalent. Noble, Boehm-Davis, and Grosz [22] have proposed a schema-based model of distributed decision making in command-and-control environments that is conceptually very similar to the RPD model already described—it is an elaboration of the recognition-matching process, using a format that has clear implications for building decision support systems that facilitate shared situation assessment.

Kahan, Worley, and Stasz [23] studied the decision making of Army commanders, and found that the commander's image of the current state of the battlefield and the desired state was more important than the generation and evaluation of alternative courses of action. Anderson [24, p. 201] reviewed historical records of the Cuban missile crisis and found little evidence for concurrent deliberation, concluding that "decision making during the missile crisis involved sequential choice over an array of noncompeting courses of action" about the relative merits of different options. Instead, the policy makers spent much of their time trying to understand the dynamics

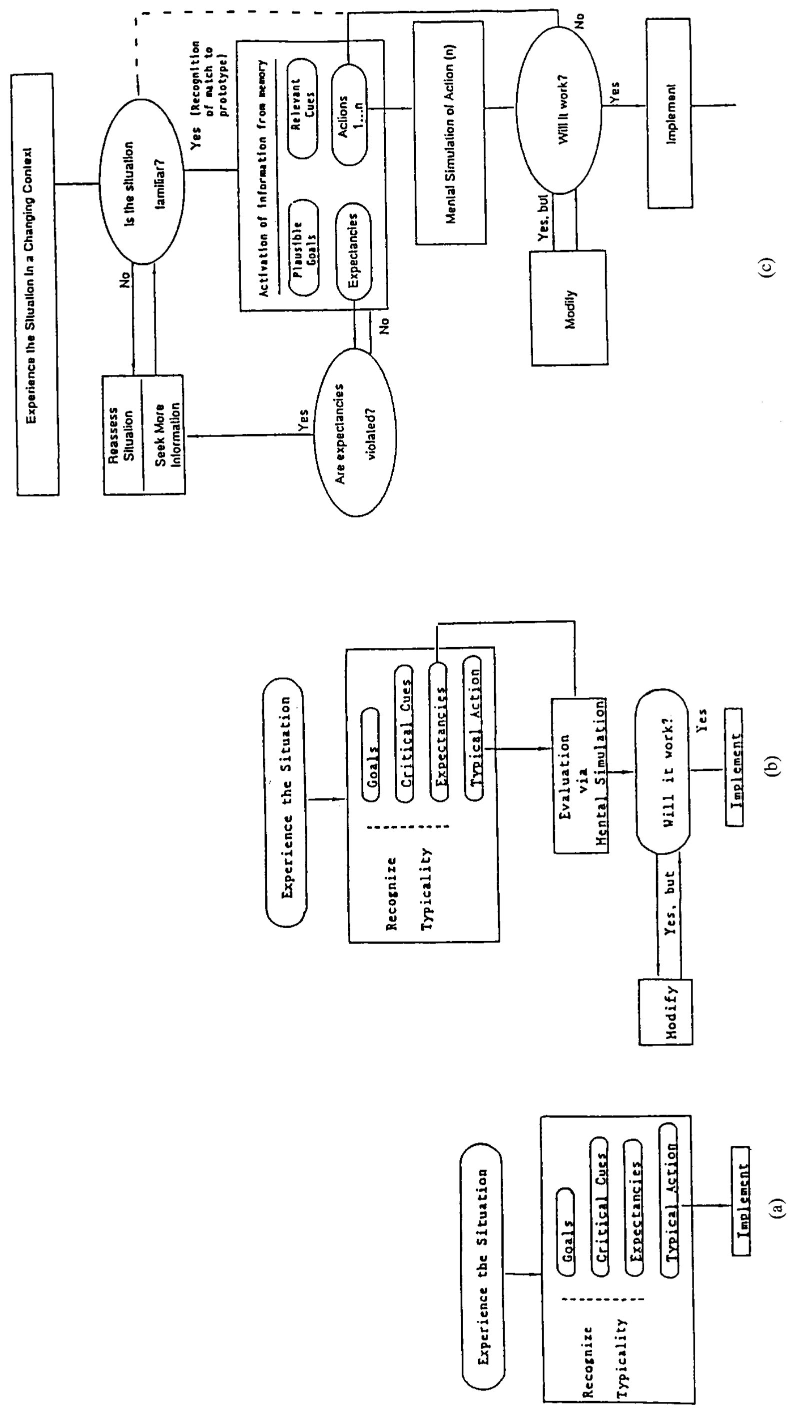

Fig. 1. Recognition-primed decision model. (a) Simple match. (b) Developing a course of action. (c) Complex RPD strategy.

of the situation and trying to anticipate how courses of action would play themselves out.

Command-and-control tasks are not the only operational environment where the decision tree metaphor doesn't fit. Studies in the business domain resulted in some of the earliest critiques of standard decision models. Soelberg [25] used protocols to establish and later test a model of the career choices of business school graduates. Soelberg argued convincingly from these data that the apparent reliance on analytic option evaluation was largely a fiction used to buttress their intuitive choice made much earlier than the business students were prepared to admit. Isenberg [26] found that business executives could recall few if any instances where they made decisions using concurrent deliberation about options. Pennington and Hastie [27] studied jury deliberation as a complex decision task and found that the major activity concerned reconstructing the events and plausible motives rather than decision making in the classical sense. Beach and Mitchell [28] have studied business decision making, and found that formal analytical models were less useful than a model of the various images used by the decision makers—images of goals, of the expected progression of the current situation, and of the impact of taking different types of action.

It is important to consider that many of the critiques to the dominant decision paradigms have been made on the basis of some form of verbal protocol analysis. Protocol analysis was a primary source for Simon's [29] now famous "satisficing" principle, and in his three-phase framework, choice was only one (and the last) phase of decision processes. Much more recently, Rasmussen [30] has used protocols and critical incident interviews to study nuclear power plant operators. His three-stage categorization of skills into sensorimotor, rule-based, and knowledge-based typology highlights some of the same task performance dimensions that we consider crucial to distinguish in models of decision making. Rule-based skill is essentially what is described in the RPD model. Considering a rule of the form: If X then do Y, a recognitional model would suggest that the expertise of the skilled operator resides primarily in recognizing when relevant antecedent conditions have occurred. One of the deceptive qualities of rules is that the consequent follows so naturally from the antecedent; it is not always readily apparent how much expertise is needed in order to recognize when the antecedent has occurred.

This is not to imply that task-oriented research is necessarily limited to protocol analysis or any particular research methods. Shanteau [31] has used a range of methods, including regression and other scaling methods, to uncover critical features of expert judgment and decision making in tasks as diverse as swine judging and weather forecasts. In a particularly thought-provoking study, Ebbesen and Konecni [32] directly contrasted the decision weights of judges setting bail for actual and simulated cases and documented a sharp divergence in the conclusions reached by this change in task. They found one set of complex decision rules when they interviewed the judges, a simpler set when they ran the judges through a simulation of a sentencing decision, and a simpler one yet (follow the recommendation of the prosecutor) when they observed the judges in action. Standard laboratory measures such as reaction time and error rates can also be used to study decision processes of nonexperts in realistic simulations of complex decision tasks, [12], [18], [33].

Hammond, Hamm, Grassia, and Pearson [34] employed a powerful design to demonstrate the importance of task parameters in inducing "analytic" versus "intuitive" decision strategies with experienced decision makers such as highway engineers. The paradigm involved changing the task characteristics to favor the use of intuitive or analytical strategies, and then varying the task information to favor intuitive or analytical decision making; the goal was to see if using a strategy compatible with the information type would enhance performance, and this hypothesis was supported. Hammond *et al.* found that for the decision of estimating traffic flow of a highway design analytical strategies were the most effective but for others, e.g., judging the aesthetic qualities of a highway design, it was the intuitive strategies that worked better. Based on these findings, they proposed a cognitive continuum between intuitive and analytical decision strategies. Klein [16] has applied this continuum concept to the difference between recognition-primed decisions and analytical decisions. The claim is that time pressure, ambiguity, uncertainty, and other such factors decrease the likelihood that a decision maker will use an analytical strategy. Rather than worrying about which strategy is better, it seems advisable to study the field conditions that favor the successful use of each type of decision strategy.

The point here is not to describe or defend any specific models of decision making, but rather to illustrate the differences in the nature of the research questions that are likely to be addressed from laboratory versus field environments. Obviously, the tree metaphor will continue to have validity for certain kinds of decision problems, but it seems quite possible that the static tree representation is most appropriate for well-structured decisions made by "novices," those who do not have the knowledge base to perform well using recognition strategies. Since it is apparently not happenstance that field studies present a more positive view of human decision performance than laboratory-based studies [35], a greater emphasis on field studies may itself go a long way in bridging the gap between research and applied needs.

IV. Decision Support Systems and Training for Recognitional Decision Making

The claim has been that there are important differences between recognitional models of decision making, which have been based primarily on studies of naturalistic tasks, and the models of behavior decision theory that emphasize some form of concurrent deliberation between options. If we can appreciate recognitional decision making as an appropriate strategy for proficient personnel to use in many operational settings, then there could be radical changes in the types of decision support systems that we attempt to design, and in the training methods we use.

One of the advantages of the decision tree metaphor is that it simplifies the analysis of a situation. It helps the researcher or practitioner concentrate on a critical aspect—the choice between action possibilities—and ignore many of the contextual factors surrounding an episode. Unfortunately, it appears that this focus does not help us to appreciate the skill of experienced decision makers, or to distinguish between good and poor decision makers, or to recommend useful decision practices. The idea that decision makers are biased suggests that training could be useful for overcoming biases (see [36] for an excellent review of decision biases). However, the concept of decision biases has been strongly criticized in recent years, raising questions about the value of this framework for training. Cohen [37] has shown that apparent Bayesian biases are not necessarily biases when examined closely. Shanteau [38] has described the limitations of decision bias approaches to applications such as training auditors. Christiansen-Szalanski [39] has shown that even statistically significant biases can have minimal practical impacts. Gigerenzer *et al.* [40] have demonstrated that biases such as representativeness may be an experimental artifact. Anderson [24] has disputed the theoretical foundation of biases such as availability, representativeness, and anchoring-and-adjustment, and Klayman and Ha [41] have demonstrated that a confirmation heuristic can improve on degrade performance, depending on the task requirements.

MAUA is a useful strategy to teach, but Zakay and Wooler [10] found that training in MAUA conferred no performance advantages for decisions that required less than a minute. Means *et al.* [11] have documented the disappointing results of programs trying to use classical and behavioral decision theory to teach general decision strategies.

A recognitional model would focus attention on the ways situations are understood, and could therefore offer a different perspective on expert/novice decision making. If system designers investigate the critical cues that operators use in categorizing or distinguishing situations, and if the designers identify the different cues used by experienced/inexperienced operators then the displays and interfaces can be centered around decisions rather than around data flows. Furthermore, if the designers can address the most difficult, time-pressured decisions then the operators should find the interfaces easier to use under stress.

Decision support systems based on a model of recognitional decision making would be designed to improve speed and accuracy of situational understanding. One way to improve recognition speed would be to use analogue displays, such as maps and graphics, rather than alphanumeric data, on the assumption that the patterns presented in analogue displays would facilitate perceptual matching. Additional facilitation could come through the use of databases organized around prior cases. There has been recent work on case-based reasoning that would be applicable here [42], [43]. Prior cases could be stored in an analogy bank so that they could be retrieved individually, or there could be a means of retrieving and synthesizing several cases at once, to allow prototype matching. The identification of one or several prior cases would allow the recognition of goals that were plausible, reactions that were typical, critical cues to monitor, and expectancies to monitor.

The verification process may be important since this is where critical errors can be caught. Decision support systems could facilitate performance by flagging mismatches or by prompting the decision maker about potentially relevant cases that were not retrieved.

The application of the prior cases for the prototype match could involve strategies for adjusting the information in the prior case: generating similar actions and modifying actions.

Turning to training, the emphasis would include sensitivity to critical factors distinguishing one prototype from another, effective use of expectancies to evaluate whether the situation has been understood correctly, and ability to anticipate the important contingencies in implementing a course of action. Training programs could address the development, verification, and modification of situational understanding, rather than teaching systematic procedures for generating and evaluating alternative options. In addition, Means *et al.* [11] have pointed out that one area where decision training appears to be effective is metacognition, teaching the decision maker to notice cues signaling information overload, confusion, and so on, and teaching procedures to better manage the decision-making task.

Rather than trying to train general skills it would be more effective to use domain-specific training programs and augment these with feedback and training scenario features that focus on decision processes, thereby leveraging the value of the existing programs. We may need to learn how to restructure existing courses so that students can learn how to conceptualize situations quickly and effectively, and to perform effective mental simulations.

V. Conclusion

What is the promise of research on naturalistic decision making? It is primarily to help us appreciate how competent people make use of their experience to handle difficult tasks under unstable conditions.

Classical decision making has examined powerful methods such as Bayesian statistics and multiattribute utility theory—methods that unfortunately are not applicable to many situations requiring support.

We can get a sense of progress by examining what we have learned about naturalistic decision strategies. In 1978, Beach and Mitchell [28] extended the work of Payne [44] and presented a contingency model of decision making that asserted that the costs and benefits associated with a task would govern whether a person used an analytical or a nonanalytical strategy. Certainly that insight seems relevant today. The analytical decision strategies Beach and Mitchell discussed correspond to the classical decision theories we have described above.

However, Beach and Mitchell ran into difficulties when they tried to describe nonanalytical decision strategies. In 1978, the best they could do was to suggest simple rules like "eeny, meeny, miney, mo," flipping a coin, or remembering homilies. Beach and Mitchell did note that habit was a nonanalytic strategy that offered some clear advantages in enabling decision makers to use experience. Therefore, we can see in this work both a limited sense of what decision makers could do other than carefully analyze the task, as well as an anticipation of some of the models presented in this paper.

Payne [45] followed this up with an article on contingency decision behavior in which scripts and production rules were mentioned as a basis for using experience to generate courses of action. However, there was no explanation of how scripts and production rules could handle nonroutine tasks, or how options would be evaluated.

Payne [44], [46] distinguished rigorous, analytical models, designated as compensatory, from noncompensatory models such as elimination-by-aspects. The interest in noncompensatory decision strategies is a useful corrective. Unfortunately, even this approach still views the more powerful methods as the ideal, and justifies the shortcuts because of the limited time and processing capacity available. Even more troubling, the noncompensatory strategies defined thus far do not include a means for decision makers to take advantage of their expertise in sizing up situations and recognizing plausible courses of action.

The decision models presented in this paper are all built on the assumption that people are reasonably competent and have strengths that allow them to handle situations even without performing rigorous analyses. These models do not view naturalistic strategies as degraded forms of ideal methods. They freely acknowledge that people will make mistakes when they use naturalistic strategies, but that these mistakes are offset by increases in flexibility and speed of reaction. We have seen a steady appreciation of this perspective over the last 15 years. The recent work on naturalistic decision making can be viewed as a current state of progress in a line of investigation into noncompensatory and nonanalytic strategies. At the same time, the work on naturalistic decision making reflects a discontinuity because of the emphasis on learning how decision makers make use of their experience.

ACKNOWLEDGMENT

We would like to acknowledge Chris Mitchell for her helpful suggestions.

REFERENCES

[1] T. Connolly, "On taking action seriously: Cognitive fixation in behavioral decision theory," in *Decision Making: An Interdisciplinary Inquiry*, D. N. Braunstein and G. R. Ungson, Eds. Kent, OH: Kent Publishing, 1982.

[2] K. Hammond, "To whom does the future belong? Is you is or is you ain't my baby?" paper presented at the meeting of the Society for Judgment and Decision Making, Minneapolis, MN, Nov. 1982.

[3] D. Meister, "The two worlds of human factors," in *Trends in Ergonomics/Human Factors*, Vol. II, R. E. Eberts and C. G. Eberts, Eds. New York: North Holland, 1985, pp. 3–11.

[4] H. Raiffa, *Decision Analysis: Introductory Lectures on Choices Under Uncertainty*. Reading, MA: Addison Wesley, 1968.

[5] D. Kahneman, P. Slovic, and A. Tversky, Eds., *Judgment Under Uncertainty: Heuristics And Biases*. Cambridge, MA: Cambridge Univ. Press, 1982.

[6] R. Nisbett, and L. Ross, "Human inference: Strategies and shortcomings of social judgment," in *New Directions in Research on Decision Makin.*, B. Brehmer, H. Jungermann, P. Lourens, and G. Sevon, Eds. Amsterdam: Elsevier Science, 1980.

[7] D. Kahneman, and T. Tversky, "Intuitive predictions: Biases and corrective procedures," *TIMS Studies in Management Sci.*, vol. 12, pp. 313–327, 1979.

[8] B. Fischoff, "Debiasing," in *Judgment Under Uncertainty: Heuristics and Biases*, D. Kahneman, P. A. Slovic, and A. Tversky, Eds. New York: Cambridge Univ. Press, 1982.

[9] J. B. Kadane, and S. Lichtenstein, "A subjectivist view of calibration," (82-6), Decision Research, Eugene, OR, 1982.

[10] D. Zakay and S. Wooler, "Time pressure, training, and decision effectiveness," *Ergonomics*, vol. 27, pp. 273–284, 1984.

[11] B. Means, E. Salas, B. Crandall and O. Jacobs, "Training decision makers for the real world," in *Decision Making in Action: Models And Methods*, G. A. Klein, J. Orasanu and R. Calderwood, Eds. Norwood, NJ: Ablex Press, in press.

[12] C. F. Gettys, "Research and theory on predecision processes," (TR 11-30-83). Norman, OK: Univ. Okla, Decision Processes Lab., 1983.

[13] J. Baron, *Thinking and Deciding*. Cambridge, MA: Cambridge Univ. Press, 1988.

[14] G. Klein, R. Calderwood and A. Clinton-Cirocco, "Rapid decision making on the fire ground," *Proc. 30th Annu. Human Factors Soc.*, vol. 1, 1986, pp. 576–580.

[15] R. Calderwood, B. W. Crandall and G. A. Klein, "Expert and novice fire ground command decisions," Tech. Rep. (MDA903-85-C-0327), Alexandria, VA: U.S. Army Res. Inst., 1987.

[16] G. A. Klein, "Recognition-primed decisions," in *Advances in Man-Machine Systems Research*, W. Rouse, Ed., vol. 5. Greenwich, CT: JAI Press, 1989, pp. 47-92.

[17] J. Taynor, B. Crandall and S. Wiggins, "The reliability of the critical decision method" Tech. Rep. (KATR-863(B)-87-07F). Yellow Springs, OH: Klein Assoc. Inc. Prepared under contract MDA903-86-C-0170 for the U.S. Army Res. Inst., Alexandria, VA, 1987.

[18] G. E. Howell, "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 Univ., 1984.

[19] W. B. Rouse, "Human problem solving performance in fault diagnosis task," *IEEE Trans. Syst., Man, Cybern.*, vol. SMC-8, pp. 258–271, 1978.

[20] J. G. Wohl, "Force management decision requirements for air force tactical command and control," *IEEE Trans. Syst., Man, Cybern.*, vol. SMC-11, pp. 618–639, 1981.

[21] R. Lipshitz, "Decision making as argument-driven action," in *Decision Making in Action: Models and Methods*, G. Klein, J. Orasanu, R. Calderwood, and C. Zsambok, Eds. Norwood, NJ: Ablex.

[22] D. Noble, D. Boehm-Davis and C. Grosz, "Schema-based model of information processing for situation assessment," (N000014-84-C-0484), Office of Naval Research, Arlington, VA, 1978.

[23] J. P. Kahan, D. R. Worley and C. Stasz, "Understanding commander's information needs" (R-3761-A), Rand Corp., 1989.

[24] P. A. Anderson, "Decision making by objection and the Cuban missile crisis." *Admin. Sci. Quart.*, vol. 28, pp. 201–222, 1983.

[25] P. O. Soelberg, "Unprogrammed decision making," *Industrial Management Rev.*, vol. 8, pp. 19–29, 1967.

[26] D. J. Isenberg, "How senior managers think," *Harvard Bus. Rev.*, pp. 80-90, Nov. 1984.

[27] N. Pennington and R. Hastie, "Evidence evaluation in complex decision making," *J. Personality and Social Psychol.*, vol. 51, no. 2, pp. 242–258, 1986.

[28] L. R. Beach and T. R. Mitchell, "Image theory: Principles, goals, and plans in decision making," *Acta Psychol.*, vol. 66, pp. 201–220, 1987.

[29] H. A. Simon, "A behavioral model of rational choice," *Quart. J. Econ.*, vol. 69, pp. 99–118, 1955.

[30] J. Rasmussen, "The role of hierarchical knowledge representation in decision making and system management," *IEEE Trans. Syst., Man, Cybern.*, vol. SMC-15, pp. 234–243, 1985.

[31] J. Shanteau, "Psychological characteristics of expert decision makers." Kansas State Univ., Manhattan, KS: Appl. Exp. Psychol. Series, 85-2. Sept. 1985.

[32] E. B. Ebbesen and V. J. Konecni, "On the external validity of decision-making research: What do we know about decisions in the real world?," in *Cognitive Processes in Choice & Decision Behavior*, T. Wallsten, Ed. Hillsdale, NJ: Erlbaum, 1980.

[33] B. Brehmer and R. Allard, *Dynamic Decision Making: A Paradigm and Some Experimental Results*. Uppsala, Sweden, 1986.

[34] K. R. Hammond, R. M. Hamm, J. Grassia and T. Pearson, "Direct comparison of the efficacy of intuitive and analytical cognition in expert judgment." *IEEE Trans. Syst., Man, Cybern.*, vol. SMC-17, pp. 753–770, 1987.

[35] J. J. J. Christensen-Szalanski and L. R. Beach, "The citation bias: Fad and fashion in the judgment and decision literature," *Amer. Psychologist*, vol. 39, pp. 75–78, 1984.

[36] A. P. Sage. "Behavioral and organizational considerations in the design of information systems and processes for planning and decision

support," *IEEE Trans. Syst., Man, Cybern.*, vol. SMC-11, pp. 640–678, 1981.

[37] M. S. Cohen, "The naturalistic basis of decision biases," in *Decision Making in Action: Models and Methods*, G. A. Klein, J. Orasanu, R. Calderwood, and C. Zsambok, Eds. Norwood, NJ: Ablex, in press.

[38] J. Shanteau, "Cognitive heuristics and biases in behavioral auditing: Review, comments, and observations," *Accounting, Organizations and Society*, vol. 14, pp. 1-2, pp. 165–177, 1989.

[39] J. J. Christensen-Szalanski, "Improving the practical utility of judgment research," in *New Directions In Research On Decision Making*, B. Brehmer, H. Jungermann, P. Lourens, and G. Sevon, Eds. Amsterdam: Elsevier, 1986.

[40] G. Gigerenzer, W. Hell and H. Blank, "Presentation and content: The use of base rates as a continuous variable," *J. Exp. Psychol.: Human Perception and Performance*, vol. 14, no. 3, 1988.

[41] J. Klayman and Y. Ha. "Confirmation, disconfirmation, and information in hypothesis testing," *Psychol. Rev.*, vol. 94, no. 2, pp. 211–228, 1987.

[42] J. Kolodner, R. Simpson and K. Sycara-Cyranski, "A process model of case-based reasoning in problem solving," in M. Kaufmann, Ed., *Proc. IJCAI-85*, 1985, pp. 284–290.

[43] E. L. Rissland and K. B. Ashley, "A case-based system for trade secrets law," in *Proc. First Int. Conf. AI and Law*, Boston, MA, 1981.

[44] J. W. Payne, "Task complexity and contingent processing in decision making: An information search and protocol analysis," *Organizational Behavior and Human Performance*, vol. 16, pp. 366-387, 1976.

[45] J. W. Payne, "Contingent decision behavior," *Psychol. Bull.*, vol. 92, pp. 382–402, 1982.

[46] J. W. Payne, J. R. Bettman and E. J. Johnson, "Adaptive strategy selection in decision making," *J. Exp. Psychol.: Learning, Memory, and Cognition*, vol. 14, no. 3, pp. 534-552, 1988.

Gary A. Klein received the B.S. degree in psychology in 1964 from the City College of New York, and the M.S. degree in physiological psychology and the Ph.D. in experimental psychology in 1967 and 1969, respectively, from the University of Pittsburgh.

He is Chairman and Chief Scientist of Klein Associates. His principal area of research is in methods of knowledge elicitation that reflect the perceptual–cognitive aspects of expertise, with applications to training, displays, and expert systems. He has performed research on critical decision making and has developed new models of decision making under conditions such as time pressure and uncertainty. Previously he had been a Research Psychologist with the Air Force Human Resources Laboratory.

Roberta Calderwood received the B.S. degree in psychology, the M.S. degree in experimental psychology and the Ph.D. in experimental psychology in 1978, 1984, and 1989, respectively, all from the University of New Mexico, Albuquerque,

She is a Human Factors Psychologist at Sciences Applications International Corporation, Albuquerque, NM. Since 1990, she has provided human factors test and evaluation support to the Air Force Operational test and Evaluation Center, Kirtland AFB, NM. Previously, she was a Senior Research Associate at Klein Associates, where she was involved in several research projects investigating decision making in operational settings and instrumental in developing decision models and knowledge elicitation methods. Her primary emphasis was on urban fireground decision making as a analogue to military command-and-control tasks.