# Rapid Decision Making on the Fire Ground: The Original Study Plus a Postscript

**Gary Klein**
**Roberta Calderwood**
**Anne Clinton-Cirocco**
Klein Associates Inc.

**ABSTRACT:** *[This is an edited version of the original, unpublished 1985 study that identified recognition-primed decision making, with a new commentary added.]* The objective of this study was to examine the way in which decisions are made by highly proficient personnel, under conditions of extreme time pressure, and in environments where the consequences of the decisions could affect lives and property. The domain of firefighting was selected, and the research focused on the decisions made by fire ground commanders (FGCs). Interviews were conducted with 26 experienced FGCs (mean experience of 23 years). Each interview covered a critical incident that was nonroutine and that demanded expertise. A total of 156 decision points were probed in this way. In less than 12% of them was there any evidence of simultaneous comparisons and relative evaluation of two or more options. In over 80% of the decision points, the strategy was for the FGCs to use their experience to directly identify the situation as typical of a standard prototype and to identify a course of action as typical for that prototype. In this way, the FGCs handled decision points without any need to consider more than one option. A recognition-primed decision (RPD) model was synthesized from these data, which emphasized the use of recognition rather than calculation or analysis for rapid decision making.

## Introduction

TACTICAL AND STRATEGIC DECISIONS MUST FREQUENTLY BE MADE UNDER EXTREME TIME pressure, yet current research in decision making has generally ignored the degree to which time pressure might influence the decision-making processes. As a means of addressing this issue, we have chosen to study the tactical decisions made at the scene of a fire by fire ground commanders (FGCs). These FGCs must allocate personnel and equipment as part of tactical planning.

We define decision making as the selection of one option from a set of two or more options. Can people make a conscious and deliberate selection of one option from a set of two or more options when they are constrained to a limited period of time? Or do people rely on other strategies that enable them to select courses of action without comparing the advantages or disadvantages of options? The study

ADDRESS CORRESPONDENCE TO: Gary Klein, MacroCognition LLC, P.O. Box 533, Yellow Springs, OH 45387, gary@macrocognition.com.

*Journal of Cognitive Engineering and Decision Making,* Volume 4, Number 3, Fall 2010, pp. 186–209.
DOI 10.1518/155534310X12844000801203. 

of FGC decision making contrasts with traditional laboratory-based methods along a number of dimensions.

### Time Pressure

Standard models of decision making postulate analytical processes that appear to be quite time consuming: identifying a full range of options, specifying evaluation dimensions, estimating utilities for each option for each evaluation dimension, obtaining scores for each option, and comparing these to determine the highest score. Surprisingly, there appears to be little data as to how the degree of time pressure might influence the use of decision strategies.

Three applicable studies (Howell, 1984; Rouse, 1978; Zakay & Wooler, 1984) were found. In both the Zakay and Wooler study and the study by Rouse, participants trained in decision-making strategies did improve performance when conditions allowed sufficient time. However, there was no evidence that the analytic strategy improved performance when decisions had to be made under time restrictions, suggesting that degree of time pressure is an important determinant of the effectiveness of decision strategies. In different experiments, Howell (1984) found that time pressure reduced participants' ability to apply their own decision rules and that time pressure combined with other variables to produce a more "intuitive" approach to problem solving, supporting the cognitive continuum theory of Hammond (Hammond, Hamm, Grassia, & Pearson, 1984). Because critical decisions made at the fire ground are frequently measured in seconds, it seemed unlikely that fire ground decisions would be characterized by the consciously deliberated processes most frequently described in the decision-making literature.

### Expertise

Many laboratory-based studies of decision making use naive decision makers in order to standardize the participants' training and experience. People are asked to make decisions about something they know little about or a problem presented in an unfamiliar way. In contrast, FGCs are experts at making the decisions that we are studying. Although each fire may present some unique challenge, the fire ground scene is their "home." Our expectation was that the experienced decision maker is quite different from the college sophomore who is grappling with a probability calculation or a move in a zero-sum game.

### Meaningfulness of Choice Consequences

Most laboratory-based studies ask participants to consider choices that have no impact on the decision maker beyond the laboratory session. At the scene of a fire, however, FGCs are making choices that affect lives and property in an emergency situation.

## Paradigm

For our data-gathering approach, we developed a method of retrospective process tracing based on the FGCs' memory of the fire scene and their step-by-step decisions and commands. A semistructured interview technique was developed for this study, loosely based on Flanagan's (1954) critical incident method, which has established the feasibility of using interview techniques for recreating nonroutine events. The general finding is that nonroutine events in the workplace are reported more accurately and completely than are more routine events. In preliminary interviews with fire chiefs we found that the most challenging incidents in an officer's career were remembered quite well. In terms of level of detail and the vividness of their accounts, this seemed to be the case for incidents occurring even 5 to 10 years ago.

We have characterized our data-gathering approach as quasi-naturalistic. We are not attempting to be "purely naturalistic" in the sense sometimes used to designate unobtrusive field observations (Brandt, 1981). We relied on interview methods wherein our participants knew they were being studied and knew the type of information we were trying to obtain. There was no deception involved. On the contrary, we were asking for their cooperation in reflecting on their decision-making skills—a fairly "unnatural" request. The naturalistic element of our approach refers to our commitment to look at decision making embedded in as much of its natural context as possible.

Quite apart from the issues relating to adequate memory for the event is the question of whether introspection is a valid means of collecting data about mental processes. Although we believe that introspection is a legitimate source of data, we do not presume that it offers a direct access to cognitive processes. It is an indirect measure with its own peculiar biases and limitations. Its attractiveness is that it offers a potentially rich source of hypotheses. The firefighters' ideas about how they make command decisions stand on their own as an important source of data. The ultimate validity in relation to any proposed cognitive model will be judged by the usual standards of scientific acceptability. Although introspective interview methods may have several weaknesses for obtaining data on mental events (Nisbett & Wilson, 1977), we felt that the possibility of capturing more of the context and phenomenological perspective of the decision maker could provide an important complement to laboratory-based descriptions.

To what extent were the methods we used gathering valid descriptions, and to what extent were the participants simply telling us what they perceived we wanted to hear? We obviously cannot determine this absolutely, but we developed a number of techniques designed to improve the accuracy and consistency of the interview data. These techniques will be described in the Method subsection on interview guide development. Our general strategy was to focus our probes in the direction of obtaining rule-based, rational calculation and option deliberation descriptions. Only when we could not obtain such a description did we probe for alternative descriptions.

# Method

## Participants

Seven fire departments from municipalities in the United States agreed to participate in the study. Some departments allowed us to contact any of their officers directly to schedule an interview; others picked certain officers to whom we could speak. In addition to allowing interviews, most departments that we contacted agreed to let us ride along on fire calls and observe fire incidents. We found our few ride-along experiences to be helpful in getting a feel for fire ground operations. However, we decided early in the study that these observations were not providing sufficient data to justify the expense of having observers on call continuously for these rare events.

### Materials and Procedure: Interview Guide Development

The interview guide was developed in an attempt to strike a balance between two disparate elicitation objectives. On the one hand, we wished the interview to be as unstructured and as free from interviewer bias as possible so that the details of the fire command would emerge with the officer's own perspective and emphasis intact. On the other hand, we did not want simply a collection of unrelated fire stories. Our perspective required that we direct the officer to focus on those elements of the incident that most affected his [or her] decision making and to structure these answers in a way that allowed the incidents to be summarized along specified dimensions.

Our solution to these conflicting goals was to ask the officer to describe the incident completely, from beginning to end, before we began our questioning. This procedure was judged to be quite successful, in part because it seemed to establish the interviewer as a listener rather than as an interrogator, and in this sense it increased cooperation. After the incident had been related, the interviewer then clarified and probed each event in the interview time line. The officer's account may have jumped around in relating the events and decisions/commands. The time line focused on representing the actual sequencing and duration of events, as well as the information and cues available at each decision point. This technique was effective for clarifying the incident events and resolving questions and inconsistencies. An additional purpose was to reactivate much of the context of the scene by asking the officer to recount the events from different time perspectives, a technique that has demonstrated utility in obtaining accurate eyewitness testimony (Geiselman, Fisher, MacKinnon, & Holland, 1985).

### Data Summary Technique

***Incident Accounts.*** The first step in the analysis was to reconstruct the account of the incident, attempting to capture in as rich detail as possible the incident from the point of view of the commanding officer. Notes and time lines were checked against the complete taped interview.

***Decision Point Structure.*** Using the completed time line and incident account, each incident was then structured into the decision format that forms the basis of the analysis. A decision point was defined as a point in time when alternative decisions or courses of action could have been chosen or taken. Thus, for each decision point there was a chosen option and one or more alternative options. This part of the analysis was largely inferential. One of the first things we learned is that the officers rarely saw themselves as either generating or selecting from a set of alternatives. We had to probe to identify options that did exist at each decision point. The FGCs experienced themselves as acting in a manner prescribed by their knowledge, perceptual cues, and goals at that moment. Thus, it was important to try to elucidate the knowledge, perspective, and cues (which we have termed "situational awareness") as they were operating and shifting throughout the incident.

The complete analysis of the decision points attempted to document the nature and chronology of the officer's situational awareness and each nontrivial decision point obtained from the incident account. Each decision point was characterized along a number of dimensions.

1. What other options were actually (or hypothetically) available to the decision maker?
2. How was the chosen option selected? Was it a deliberated choice? Could a selection rule be articulated or inferred?
3. How much time was taken in making the decision?
4. How much time pressure was involved in making the decision?
5. What level of experience was required to make the decision? How much experience was required to interpret the cues or know which cues to look for? Was there a rule that could be implemented by a less experienced officer? What kinds of critical knowledge or cues were found to be missing?

Questions about time and degree of time pressure were difficult for FGCs to estimate. In many cases the answer to the time probe was simply "immediately" or "I just did it automatically, based on experience." These expressions were coded as less than 1 min, although in verifying this with the participants, they indicated that it was actually fewer than 30 s. Time pressure was also difficult for the officers to report, so it was largely inferred on the basis of the time line information.

## Quality Control Procedures

Early in the course of conducting interviews, we found that it was difficult for one person to capture all of the relevant aspects of the interview, including obtaining and probing a complete time line. It was therefore decided that it was preferable for two interviewers to be present. However, for seven of the interviews there was only one interviewer present. Also, because interviews were generally 2 hr long, it was decided to tape each interview so that later discrepancies and questions might be resolved.

Coding was a lengthy process, requiring anywhere from 3 to 7 hr per incident. Because of the number of incidents we wished to obtain, it was not feasible to implement any formal procedures for obtaining an assessment of intercoder reliability. The following informal quality control procedures were employed: (a) Each incident account and decision analysis was read and criticized by the other member of the interview team or by one of the other interviewers, and (b) each incident account and decision analysis was then reread and questioned by each of the two principal investigators (who also may have been a member of the interview team).

## Results

### Incident Characteristics

We collected 32 critical incidents. They were collected from a total of 29 interviews that were conducted with 26 officers. In 3 cases, the same officer was interviewed twice concerning different incidents. In 2 other cases, multiple incidents were recounted in a single interview session because the officer's account of the initially selected incident was extremely short, yielding only a few decisions. [The 29 interviews generated 29 separate incidents but resulted in a total of 32 critical incidents because of the following reasons.] One fire, at an oil-pumping station, was so large that it was counted as 3 separate incidents (roughly covering 3 separate days of the incident) and was recounted by 2 different officers. Two incidents were separate versions of the same fire given by officers of different rank, offering different perspectives. Of the 32 incidents, 29 were fires, 2 were rescue operations, and 1 was a gas leak.

The officers interviewed were of high rank and experience: 6 lieutenants, 4 captains, and 16 chiefs. The 26 officers had an average of 23.2 years of firefighting experience. None of the interviewees had less than 12 years of experience, and the maximum was 37 years. The interviewees also had command experience. For 28 of the incidents, the interviewee was the FGC initially in charge at the scene. The other 4 incidents involved officers in charge of squadrons or sectors.

The criterion for selecting a particular incident was that it presented a command challenge or was nonroutine in some way. We found that the officers' reasons for selecting an incident could be characterized by four factors. Any of the factors that applied were checked for each incident. Fourteen were designated as recent incidents, 13 were of unusually high risk, 5 contained disappointments in terms of the outcome, and 12 were designated as primarily nonroutine. [Twelve incidents were not coded for risk.]

In general, the selection of incidents did provide us with a good range of cases for study. There were rescues; fires in single residences, apartment houses, hotels, businesses, and factories; and an oil tank leak amid a large complex of oil tanks. We were concerned with the experienced FGCs' decision process and not the correctness of the decisions. In some cases we found errors in judgment, and in the case of the oil pumping station, the situation was so unique and complex that the FGC had little experience in handling the situation and eventually needed to bring in consultants for guidance. A few of the incidents were selected because they involved dramatic search and rescue operations. These were vivid for the firefighters but tended to be less interesting from a decision-making perspective.

The incidents selected generally occurred within the year prior to the interview. The median recency was less than 1 year, and this was the mode as well. More than a third of the incidents had occurred less than 3 months prior to the interview. Four of the incidents had occurred more than 5 years earlier.

Officers were also asked to rate the incident on four separate risk factors: the risk to the initially burning structure, the risk to the adjoining structures or property, the risk to civilian life, and the risk to firefighters' lives. Ratings used a 3-point scale with 1 indicating *low risk,* 2 *medium risk,* and 3 *high risk.* All the incidents were classified as high risk on at least one dimension. The mean ratings for the 29 unique incidents were 2.7 for risk to the involved structures, 1.6 for the risk to adjoining property, 2.2 for the risk to civilian 1ife, and 2.6 for risk to firefighters' lives.

Containment is a point at which the fire is no longer escalating, not including the final fire control operations or salvage. The median reported time to contain a fire was 2 hr. The shortest fires were contained in 15 min, and the longest lasted 12 hr, with the exception of the pumping station fire, which took 1 week to contain. The total alarms called for a fire gives a rough estimate of seriousness, although procedures vary widely from department to department. Seven of the incidents were 1-alarm calls, 6 were 2 alarms, 8 were 3 alarms, and 2 were 4 alarms. [Six of the incidents were not coded for number of alarms.]

## Analysis of Decision Points

The basic unit of analysis in the study was the decision point, the point in time where multiple options existed. For the 32 incidents studied, we analyzed a total of 156 decision points, with an average of 5 decision points per incident. The number of decision points ranged from 1 to 10 per incident.

***Time.*** As expected, the decisions were generally very time critical; 122 of the 156 decisions (78%) were reported to have been made in less than 1 min, with many of these being made in under 30 s. Another 15 were estimated to have taken 1 to 2 min. Ten took 2 to 5 min. Nine of the decisions took more than 5 min, but all of these were from the pumping station incident, which was a unique case. First, it took almost a week to contain, which is an order of magnitude longer than even the longest of our other cases; second, the local firefighters who tried to contain the blaze could not be considered experts. The scope of the fire and the problems encountered were outside their experience. In addition, there were problems in coordinating the efforts of the several departments involved, so there was initially no clearly designated FGC. Thus, in the pumping station fire, many of the decisions were made in consultation and were stretched out over several hours.

***Time Pressure.*** The degree of time pressure is conceptually distinct from how rapidly the decision was made. A decision might be made very rapidly simply because it can be, not because it must be. We rated each decision point for its degree of time pressure using a 4-point scale in which 1 = *low time pressure* (incident was stable), 2 = *some potential for escalation,* 3 = *imminent loss of control,* and 4 = *threatened loss of life.* A majority of the total decisions were made under conditions of some extreme urgency; 61% of the decisions were ranked as either time pressure Level 3 or 4 (95/156). For these levels, every second was important for combating the exponential increase in the fire. Another 36 fell at urgency Level 2, in which perhaps minutes were available for making the decision. Only 13 were made under low time pressure. All but two of these decisions were from the pumping station incident. [Twelve incidents were not coded for perceived time pressure.]

***Decision Type.*** One of the most important goals of this study was to describe how the officers reported making their decisions. We had originally expected that the FGCs would tell us about the decisions that they had trouble making. This rarely happened. In almost none of the cases did an FGC even report making a decision in terms of comparing two or more options and trying to select one. In other words, we found virtually no instances of the standard laboratory paradigm for decision making: conscious and deliberate selection of one alternative from several. We tried several coding schemes to describe the way that the FGC handled each decision point. Table 1 shows the type of decision strategy found for each of the 156 decision points.

One category we considered was option selection, whereby the FGC would receive the options from an external source and select one. This category does describe a standard way that decision making is studied in laboratories. However, we saw no evidence that any of the FGCs used this approach. In fact, we found only a very small subset of decisions for which alternative options were even considered, let alone concurrently contrasted. For only 10 of our decision points, not including the pumping station incidents, we found that the FGCs consciously contrasted options in arriving at a decision ("deliberated"). In these cases, the FGC would typically identify two or more ways of accomplishing a goal and then would make the selection on the basis of a single dimension, or only a few dimensions, such as the time required and risk factor involved. There was never any systematic examination of all the relevant attributes in some attempt to perform a rudimentary decision analysis. Another type of decision was when the FGC faced an unfamiliar situation and had to creatively generate or construct the possible options ("constructed"). There were 11 cases that fell into this category.

In the case of the pumping station, roughly 59% of the decisions were arrived at through a consciously deliberated process, almost always in group consultation. Therefore, the deliberation component may apply more to the group aspects of the decisions being made than to decision making for any individual. The majority of the decisions were characterized not by option consideration but by the FGCs recognizing the situation as an example of something they had encountered many times before. In other words, there was evidence for a matching process rather than a calculational process.

TABLE 1. Frequency Distribution of Decision Strategy Used for Each Decision Point

| Type | Frequency | |
|---|---|---|
| Option selection | 0 | |
| Deliberated | 10 | (8) |
| Constructed | 7 | (4) |
| Procedural rule | 0 | |
| Analog | 3 | |
| Prototype | 114 | (10) |
| Total | 134 | (22) |

*Note.* Pumping station data in parentheses.

The first type of matching we looked for was matching to a specific analog—another situation they had been through or heard about. We found very few cases of this. We specifically probed for analogs and found only three. There were no cases in which one fire was treated as an analog for another. Rather, the analogs seemed to serve as flags, which alerted the firefighter to dynamics of the situation that needed attention. Their effect was on situational awareness and on specific decision points. In Case 5, for example, the FGC looked up at a billboard near the roof of an apartment building that was burning down and remembered another time when a billboard had collapsed, falling off the roof and posing a hazard to firefighters and civilian onlookers. He therefore ordered the crowds to be moved back.

Apparently, the FGCs had so many similar firefighting experiences that these became merged in memory, with no specific cases standing out. A fire could be spoken of as typical, which suggested our next category: *prototype*. For example, they have been through hundreds of house fires. When they encounter one, they view it as typical of their prototype, which would include some standard layouts, some standard approaches, and so on. We used the notion of a prototype in a way that overlaps the concepts of scripts and frames, and we will discuss the theoretical issues further in the Discussion section.

It is important to clarify how we used this category. The FGCs encounter a decision point, they recognize a match to a prototype, and the prototypical scenario guided by experience tells them how to proceed. In this way, they implement a course of action without ever considering any of the other options at the decision point. In our interviews, we probed this very carefully, and the FGCs were clear that they were not aware of other options. That is why they did not feel that they were making decisions. To ensure that these were really decision points, we probed to identify potential options. Usually, the participants were not able to find any. In these cases, we asked about options that a novice might be foolish enough to consider. Sometimes we had to suggest the options. As long as we found at least two options, we considered it a decision point. We did not study how many options existed, given that we were now dealing with the hypothetical.

For the 156 decision points, 127 fell into the prototype category. This was the dominant approach. Options were selected without any reports of conscious examination, evaluation, or analysis. In most cases, the events triggered an immediate cognition of what had to be done, and the action was taken. However, there were three decision points at which there was not a match but a lack of an expected match, and this mismatch triggered a new situational awareness and the reassessment or shift to another plan.

In Case 4, a firefighter led his men into a burning house, round back to the apparent seat of the fire in the rear of the house, and directed a stream of water on it. The water did not have the expected effect, so he backed off and then hit it again. At the same time, he began to notice that it was getting intensely hot and very quiet. He stated that he had no idea what was going on, but he suddenly ordered his crew to evacuate the house. Within a minute after they evacuated, the floor collapsed. It turned out that the fire had been in the basement. He had never expected this. This was why his stream of water was ineffective, and it was why the house could become hot and quiet at the same time. He attributed his decision to a "sixth sense." We would be less poetic and infer that the mismatch was the cue. The pattern of cues deviated from the prototypical patterns in which heat, sound, and water are correlated.

Another category we tried to use was a *procedural rule*. In this case, there would be a rule of the form "If *x*, then do *y*." All the FGC would have to do is match the *x* condition in order to determine whether to implement the *y* action. In practice, we had difficulty in distinguishing this from matching to a prototype. In both cases there is a matching, followed by an action. The only difference is that the rule is context free, whereas the prototype is context rich. We were not able to find any examples of context-free rules that could safely be initiated by novices. If the concept of a procedural rule is broadened to include contextual sensitivity, then the border between procedural rule and prototype matching becomes very blurred.

Table 2 is a summary of decision time and time pressure findings for each decision type. It simply confirms the tendency for deliberated and constructed decisions to be made under less time pressure and to take longer to make than decisions made on the basis of a prototype match or mismatch.

***Situational Awareness.*** Because we treated decision making as a form of complex pattern matching, much of the expertise of the FGCs came through in the situational awareness. This reflected their understanding of the dynamics of the case and was the basis for their ability to recognize cases as examples of standard prototypes. In many of the cases, the initial situational awareness was maintained throughout the incident, with new information serving to elaborate on what was originally known.

In other cases, there were dramatic shifts in the situational awareness. For example, in Case 23, a fire at a chemical plant, the situational awareness included the dynamics of a burning structure and flowing chemicals as well as the risk of nearby storage tanks exploding. Foam is the first choice for putting out a chemical fire; it extinguishes the fire by smothering it. However, the tanks required cooling to reduce the risk of explosion, and water is the best coolant. A novice may have used foam initially to try to extinguish the fire, or used both water and foam, which would have produced a diluted and ineffective foam.

TABLE 2. Frequency Distribution of Decision Types, Time, and Time Pressure

| Type | *N* | Time (min) | | | | Pressure Level | | | |
|---|---|---|---|---|---|---|---|---|---|
| | | *<1* | *1–2* | *2–5* | *>5* | *1* | *2* | *3* | *4* |
| Prototype | 117 (10) | 108 (6) | 7 | 2 (2) | 0 (2) | 2 (1) | 33 (7) | 47 | 35 (2) |
| Deliberated | 10 (8) | 5 | 4 (1) | 1 (2) | 0 (5) | 0 (1) | 5 (7) | 5 | 0 |
| Constructed | 7 (4) | 3 | 3 | 1 (2) | 0 (2) | 0 (1) | 2 (2) | 3 (1) | 2 |
| Total | 134 (22) | 116 (6) | 14 (1) | 4 (6) | 0 (9) | 2 (3) | 40 (16) | 55 (1) | 37 (2) |

*Note.* Pumping station data in parentheses.

The expert's decision was to use water initially to cool the tanks and then to shut down the water and apply foam. He used his perceptual ability to judge when the tanks had been appropriately cooled, so that an explosion was no longer likely. He relied on such cues as heat waves and steam levels coming off the tanks. When the foam operation was begun it was monitored and judged effective. Later, however, runoff was discovered to have been seeping into the basement of an adjoining structure, creating a new fire hazard. The situational awareness was now changed, and the FGC recognized a new out-of-control situation. He immediately called for additional personnel and equipment to handle the expanded situation.

This example shows the perceptual ability needed, the ability to rapidly assess the situation, the ability to shift this awareness, and the ease of making decisions. In our data analysis we decided not to define a change in situational awareness every time anything happened or failed to happen. This would have become cumbersome. We reserved our changes for those events in which there was a shift in goals or subgoals as a result of new information. For most incidents, there were generally 3 to 5 different situational awareness changes. Rarely did we identify more than 10.

In our coding for situational awareness we developed a checklist of nine dimensions that seemed useful. These are presented in Table 3. The exact number of dimensions or the ones selected are not important here. What is important is that these dimensions reflected different classes of causal factors that were being learned and interpreted by the FGCs to suggest and constrain courses of action.

In determining how situational awareness was growing, we had to infer what expertise was needed to interpret new facts and to perceive changes. The elaboration of the required expertise took the form of a knowledge analysis, or a critical cue analysis. It described the type of knowledge and recognitional ability that the FGCs needed to handle these critical incidents. As such, it was quite different from the standard firefighting procedures offered as guidance. Instead of vague statements about how the FGC needs to be able to determine when water is hitting the seat of a fire, we have specific cases, such as Case 21, in which the FGC thought he was getting at the seat of the fire, waited for signs of white smoke that show a fire is being extinguished, saw none after about 30 s, and began to worry that he did not know where the seat of the fire was at all. This type of description defines the cue, the nature of the cue changes being assessed, and the time frame for expected cue changes.

## Discussion

The study was successful in challenging some of our basic assumptions about decision making and forcing us to reconceptualize our approach. In this section, we will discuss several aspects of that reconceptualization: (a) a recognition-primed decision model, (b) a characterization of situational awareness, (c) methods for describing perceptual learning, (d) related processes such as analogical reasoning and imagery, and (e) types of deliberated decision making. Finally, we will discuss

TABLE 3. Situational Awareness Cue Checklist

1. Problem
   - Smoke: color, amount, toxicity
   - Fire: amount, location
   - Explosion potential
   - Chemicals
   - Rate of change
2. Structure
   - Type: factory, house, office, vehicle, etc.
   - Materials: wood, brick
   - Architecture: special features
   - Age
3. Problem × Structure
   - Seat of fire
   - Possibilities for movement
4. Weather
   - Temperature
   - Moisture
   - Wind: velocity, direction
5. Risk to life
   - Direct cues
   - Knowledge of potential risk
   - Special populations: older adults, disabled, etc.
6. Risk to firefighters
7. Nature of attack
   - Progress
   - Hindrances
8. Resources
   - What is available?
   - What is needed?
   - Special needs
9. Goals assessment
   - Search and rescue
   - Fire control
   - Property conservation

the accomplishment of the project objectives and some of the implications of this research for issues such as decision aids and the selection and training of decision makers.

## Recognition-Primed Decisions

The major finding of this study was that FGCs rarely reported having considered more than one option. In an analysis of 156 decision points, we found that in only 28 was more than one option even identified. In only 16 did the FGC report doing any relative evaluation of one option versus another, and these were for cases specifically selected for their difficulty. If there were instances in which they would have had to wrestle with choices, it would have been these cases. We therefore concluded that the standard approaches to decision making would not apply to the vast majority of these cases.

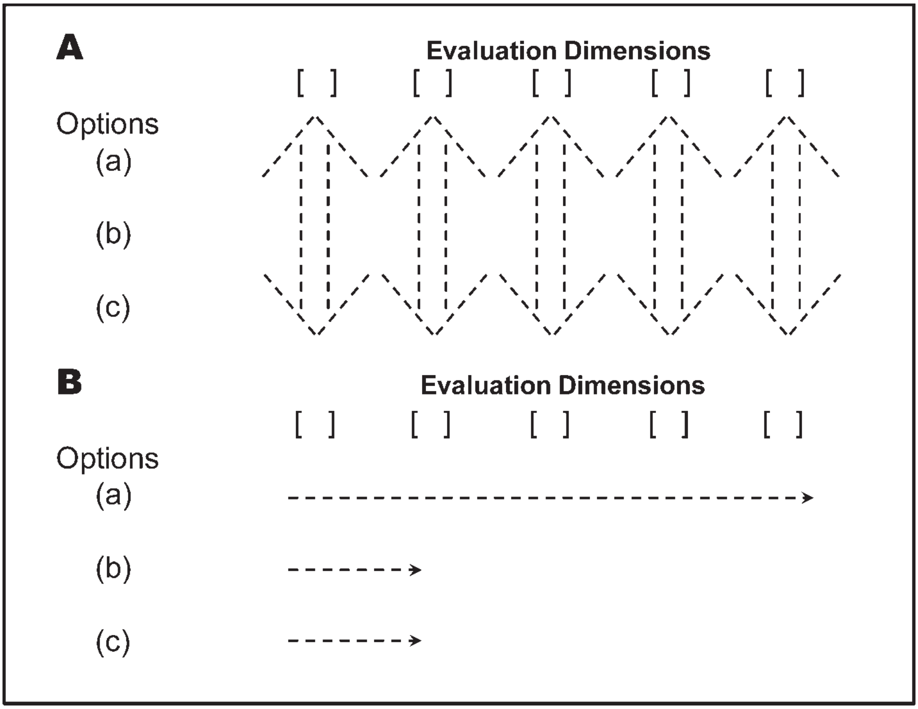


*Figure 1. (A) Concurrent evaluation: vertical model. (B) Sequential evaluation: horizontal model.*

Their ability to handle decision points depended on their skill at recognizing situations as typical instances of general prototypes that they had developed through experience. The prototypes provided them with an understanding of the causal dynamics at work, suggested promising courses of action, and provided them with expectations.

By contrasting concurrent with serial models of option evaluation, the FGCs' decision-making strategy can be more accurately described. Figure lA shows a standard decision analytical structure, with the options down the side and the evaluation dimensions across the top. In such a model, the decision maker is presumed to consider several options at the same time, perhaps performing pairwise comparisons, and to make concurrent, conscious judgments and reflective evaluations of the strengths and weaknesses of the different options. In contrast, Figure lB shows a serial option evaluation model. Here, an option is generated and then either implemented or rejected. If rejected, a second option is considered, and so forth. This may be described as a serial model of decision making, because although one or more options is considered, only one option is examined at a time. We feel that this model is a better fit for the data we collected.

We were specifically studying decision making under extreme time pressure, when there is a need for rapid decisions. We therefore characterize our description as a *recognition-primed decision* (RPD) model. We have developed a different approach to decision making in part because of the way we conducted this study. We looked at decisions made under extreme time pressure performed by personnel with high levels of experience and personal commitment. The use of the quasi-naturalistic design has given us a unique perspective. Standard laboratory studies usually do not use highly experienced participants. These studies tend to focus on the analytical skills needed to evaluate options, and they leave option generation as something of a mystery. Campbell (1960) described this sort of approach as random generation and selective retention: basically a random generation of options, followed by analytical methods to identify and select the best option.

For our participants, option generation fell out of their situational awareness. They were able to identify good options immediately: This was part of their skill. Von Clausewitz (1832/1976) referred to this ability as *coup d'oeil*, the skill in making a quick assessment of a situation and its requirements. In contrast to Campbell (1960), we are placing more of the burden on the recognitional processes and the use of experience to generate a plausible option as the first one considered.

For a time-limited task, concurrent evaluation is probably impossible. It takes a lot of time to consider all options along all evaluation dimensions. Even reducing the number of options and dimensions still places an unreasonable load on the decision maker. Serial evaluation seems necessary. Decision researchers may have to study situational awareness and expertise to understand how effective options are immediately recognized.

The advantage of the RPD model is that it provides the decision maker with a course of action at every point. The decision maker begins with an initial option, and if a response is called for, this will be 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. In the situation we studied, it is critical that the decision maker always be ready to act effectively. The RPD model may explain some aspects of intuitive decision making (Hammond et al., 1984). If people are using recognitional and perceptual matching processes, it would be understandable that they could not articulate the bases for the decisions. Our RPD model is also consistent with Simon's (1955) notion of satisficing. Simon introduced the concept of satisficing to point out that decision makers typically do not do all the work to ensure optimal decisions. He was criticizing the economic assumption of the "rational man." However, he did not offer an explanation for how satisficing could be accomplished. How can people be sure of finding workable options without generating all options possible and evaluating these options?

Several approaches to serial evaluation models have been proposed. A lexicographic method (e.g., Aumann, 1964; MacCrimmon, 1968) postulates that a decision maker generates a set of options, orders the evaluation dimensions, starts with the most important, and selects the option that has the highest rank on the most important dimension. If there is no clear winner, the second dimension is examined, and so on. This model allows serial evaluation but still concurrent consideration of options, and it does not fit the majority of our data.

A second model is Tversky's (1972) elimination by aspects. Here, several options are generated, they are evaluated on the first dimension, any that do not pass a criterion are rejected, the remainder are evaluated on the second dimension, and so on. Again, this is concurrent consideration of options and serial evaluation on dimensions. This model does not fit our data because it stops when one option is left remaining. This could still take a long time, and it is unrealistic because an option might still be unacceptable on remaining dimensions. Finally, it does not allow the generation of new options without starting the whole process over. Clearly, this will not do for a time-pressured task.

A third model is that of conjunctive standards (Bettman, 1971; Kleinmuntz, 1968). Here, there are criteria for each evaluation dimension. One option is generated, and if it fails to meet any of the criteria it is rejected and a second option is evaluated. This is a better description of the data we collected, but there are still a few shortcomings. This model lacks a description of how effective options can be generated immediately, something we will discuss later in this article. More important, our analyses suggested that options are not evaluated as systematically as in a model of conjunctive standards. We did not see FGCs carefully examining any option to see if it passed the criterion point on several dimensions. Rather, there was a general matching of the current situation to the prototype. If the match was good, then the option was implemented. If it was questionable because of differences in causal dynamics, then the FGC began to consciously evaluate, often by imagining how the option would be carried out and what difficulties might arise. In some cases, the option could be augmented to handle the difficulties. In other cases, it had to be rejected.

This approach is basically the one described by de Groot (1946/1978), who studied the way grand master chess players would select a favorite option and explore its adequacy through progressive deepening. The grand masters were not looking at all options and conducting shallow evaluations (the way a computer program usually does).

An example might be helpful. In Case 9 there was a call to rescue a semiconscious woman who had jumped off a highway overpass and was dangling suspended on one of the metal supports for a sign. The FGC rushed to the scene. Two of his crew had climbed onto the supports and were holding her arms and legs. In that situation, the immediate need was to provide a firm basis of support for her. The standard approach is to use a Kingsley harness, which snaps onto a victim quickly and allows the victim to be moved and raised. However, the woman's position was not standard—that is, she was face down, and a Kingsley harness is strapped on from the front.

The FGC imagined moving her into position and realized this would place her and his crew in danger. He also felt that attaching it from the back, a nonstandard approach, would create severe strains on her back. (Tests the next day showed he was right.) He rejected this option. Next, he considered another standard type of rescue equipment, a Howd strap. This loops onto a victim in different ways, but again, the match wasn't right. Howd straps are also attached from the front and are open to the same weaknesses as the Kingsley harness. He rejected this option. Next, he thought about ways to use the Howd strap differently but could not come up with any strategies simple enough to guide his crew through. Finally, he remembered the ladder belt, which firefighters strap around their waists and clip to their ladders to make sure they will not fall off during a rescue. Ladder belts can easily be attached from the back and have only one buckle. Moreover, a simple rope could be used to attach the ladder belt to the crew above her. He quickly ordered the ladder belt brought out. This decision took less than 1 min from the time he first arrived at the overpass.

This example shows how the FGC examined four options, all serially. The standard method was considered, evaluated, and rejected; another standard method was considered, evaluated, and rejected. Next, an attempt was made to construct an option, and then when this seemed unlikely, a second constructed option was identified, one that had never been used before in that type of situation. There was never any attempt to compare two options at the same time.

## Situational Awareness

In the RPD model of time-pressured decision making, situational awareness becomes very important. We are claiming that most decision points can be handled without deliberation by applying an "if *x*, then *y*" strategy. The key to making this work is that an effective option be immediately identified in the majority of cases, and the way that experienced FGCs can identify effective options is to match the current situation to a prototype, thereby recognizing it as typical and amenable to typical procedures. The recognition of a situation as typical of a prototype depends on the way the FGC has assessed the scene and its problems—that is, the ability to know that "*x* applies" is dependent on situational awareness. The situational awareness provides the information for triggering the conditional.

For example, in Case 12, a simple residential fire, the FGC sent a crew into the building with a hose to hit the seat of the fire. The rule might have been "if there is a fire in a house, hit the seat of it." It is a standard procedure, apparently requiring little expertise. However, upon probing we found that another procedure might have been to send the hose around the house, break some windows, and hit the fire from the back. This might even have been faster. The FGC never considered it. The basic idea is not only to hit the seat of the fire but also to drive it out of the house. Hitting the fire from the back would only drive it farther into the house. Continuing our probing, we asked if anyone would have sent the hose around back. The FGC said that this was done too often by people who should know better, trying to reduce the risks to their crews. It might be appropriate if there was no one in the house, or if the house was not worth saving, or if there were adjoining structures that could be endangered by an internal strategy. All of a sudden, a simple decision became complicated.

Part of the skill of an FGC is in knowing when to obtain more information. The term they use is "size-up." Theoretically, a size-up must be done at every incident prior to acting. However, for practical reasons, it is not always possible to complete the size-up. Valuable time can be lost continuing a size-up while a fire spreads out of control. For example, in Case 26, a fire in a factory, the FGC began to walk around the factory, saw a fire burning through the wall, and immediately ordered his crew to train their hoses on it. He recognized that it was a good place to use his resources and did not want to let it burn further while he kept walking. His experience allowed him to judge what a good place to hit a fire looked like. A novice would not be able to make such a judgment. In this domain, time is critical and actions must be initiated without the benefit of complete information analyses.

## Perceptual Learning

The FGCs showed an impressive variety of perceptual learning. For example, Case 21, a fire in a plastics factory, required that the FGC interpret the smoke color, the color of the fire, and the sponginess of the roof to assess the situation. At first, there was only smoke coming out of the front of a factory. The FGC assessed this as a simple fire and trained his hoses on the source of the smoke. If he had been hitting the seat of the fire, the smoke should have turned white within 30 to 60 s. It did not, and he concluded that the fire was burning farther inside.

He then sent some firefighters to the roof to open a channel that would let the heat and smoke escape. They reported that the roof in a back section of the factory felt spongy. He went up to investigate because, to inexperienced personnel, all roofs can feel spongy. He found that the roof indeed had a spongy feel, and ordered his crew off of it. He concluded that the fire was larger than he had thought and was probably burning directly below them. There is no way to describe what a spongy roof feels like. This recognition comes only with experience of walking on roofs that are solid and roofs that are spongy and learning to discriminate between them. Finally, they found the seat of the fire. Its bright orange color suggested that the heat was in excess of 1,000° F (538° C), and a second alarm was called.

The interview guide approach that we used highlighted the perceptual cues used by the FGCs, and our coding system retained this information by linking these critical cues to situational awareness and to RPDs. Therefore, our approach may be useful for developing a critical cue inventory of the types of cue discriminations that must be maintained for expert performance. It may have value as a knowledge elicitation method that would feature perceptual knowledge-based rather than rule-based knowledge.

The RPD model is summarized in Figure 2. The current situation matches a prototype based on similarity of goals, perceptual cues at the scene of the fire, and causal factors and information about these. The prototype generates expectancies and also a set of options, with the most typical option generated first. The action is evaluated for plausibility and is implemented, modified, or rejected. Often if there is an unfamiliar situation, the evaluation will include imagery of the anticipated consequences of using that option. If the option is rejected, the next most available, representative, and similar one is selected for evaluation.

## Related Phenomena

We found several phenomena during this study that may be of interest. One was the use of imagery. In evaluating options, a common strategy was for the FGC to

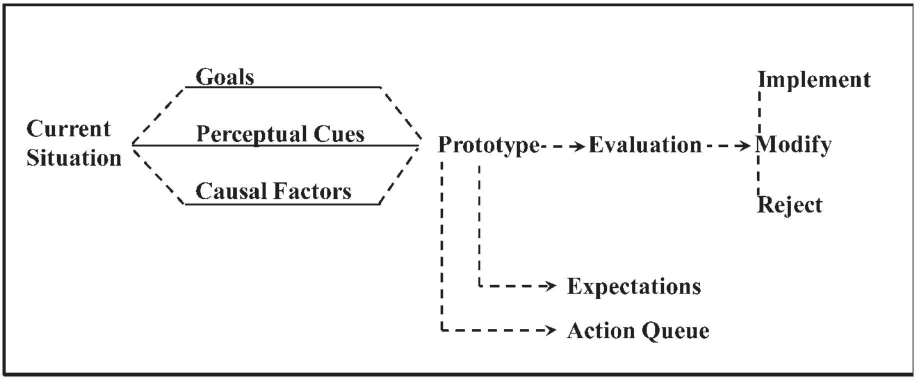


*Figure 2. The recognition-primed decision model.*

create an image of how the option would be implemented. This strategy was used rather than the analytical method of evaluating the option with regard to criteria on several dimensions. A second phenomenon of interest was the failure to find evidence of analogical reasoning. We had originally expected to find a great deal of analogical reasoning. We found little. The reason is probably that with increasing numbers of similar experiences, analogs become less vivid and fade out altogether, becoming prototypes. We expect that for less experienced personnel, analogs are still important. Where we did find analogs, they were to specific cues and features of a situation, especially nonstandard ones.

It should also be noted that we are using *prototype* here to refer to syntheses of analogs. We are not claiming that episodic memory has transformed into semantic memory (Tulving, 1972), as this would involve the degeneration of context-rich episodes into context-poor semantic networks of abstract elements. We think this is wrong because the FGCs were still very sensitive to context, especially as it alerted them to features of situations that had potential implications. The way we are using *prototypes* includes a highlighting of the general features of situations along with a potential for specific details where relevant. A third phenomenon is the reliance on constructed options—cases in which the “if *x*” portion of the RPD has been satisfied and the means of implementing the “then *y*” portion needs to be found. Usually there is a standard means, but sometimes the FGC will need to find a unique means or will reject the standard approach in favor of an innovative one. This is important because decision-making models must be able to describe how people can synthesize options to develop new and improved options. If decision making is treated entirely as a way of selecting between a fixed set of options, then the opportunity for growing new options will be missed.

## Types of Deliberated Decision Making

Although most of the decision points were best described by an RPD model, about 10% (not including the pumping station incidents) included some sort of conscious processing for selecting one option from several that were considered.

In none of these cases did the FGC attempt to identify additional options. The focus was on the primary options that were identified. In none of these cases did the FGC go through an exhaustive evaluation of relative advantages and disadvantages on each important evaluation dimension. Once the FGC was confident that each option was worth considering, the decision was reported to be made on the basis of a single overriding dimension, such as safety or time savings.

One fire does stand out as a contrast to our general findings, and this was the subject of three interviews, Cases 30, 31, and 32. It was the fire at a pumping station.

The command structure and experience at this fire were unlike those at our other incidents. Fire companies from six small towns were called in and somehow had to be coordinated on the spot. The cold forced most firefighters to wear masks for warmth, which disrupted the chain of command because the FGC was not easily identified. Basically, the fire chiefs were reduced to the level of novices, given that they were working with something outside their usual experience. They had no prototypes for action, and they did not take many actions. Success was largely attributable to their decision to bring in consultants who did have experience with fires of this magnitude and who were immediately able to see what was needed and what courses of action to take. The pattern of their decision making does conform more closely to the standard accounts. They consciously worked to identify options. They were unsure about the advantages and disadvantages of options, and they had to try to figure these out analytically because they did not know what to expect.

## Conclusions

We intended to assess the value of a heuristic model of decision making. This was accomplished, but the model we were assessing was found to be inadequate. We had hypothesized that FGCs did not have time to consider a wide variety of options, and we speculated that they would use a strategy of considering only two options at a time, for purposes of direct comparison. We thought that this was a radical hypothesis, but in fact it turned out to be too conservative. In general, they only considered one option at a time, as described by the RPD model.

We wanted to increase our understanding of decision making by experts under high time stress. We were partially successful in this. We have proposed a descriptive model, the RPD model, and we have a better appreciation of the role of situational awareness. Yet we realize that there is still much to learn. In hindsight, our estimate of the time it takes to make decisions was overly focused on the time to choose the course of action. Our time estimates do not include the "time to situational awareness," which in many cases is the most important aspect of the decision.

One concern is with the validity of our findings. Verbal protocols as a data source have a long history in psychology, but have at various times they have been viewed as an invalid variant of introspection. In our study, we tried to avoid the tendency for participants to speculate by asking them to recreate as much as possible what they were actually seeing, hearing, and thinking at a specific moment. We never simply asked "why" an action was taken, which is a procedure Nisbett and Wilson (1977) have criticized. Furthermore, we suspect that asking experienced personnel to reconstruct aspects of their decision processes is different from asking participants to speculate on their motivations in an unfamiliar laboratory environment. In our case, even when our probes directed officers to relate possible options that they might have considered, the FGCs firmly insisted that this was not what they actually did. Thus, although their recall may still be inaccurate, it is not likely attributable to acquiescence to experimenter bias. Regardless, we do not present our data as firm evidence supporting our model. Rather, we are suggesting that the retrospective protocols enabled us to generate new hypotheses.

We were also concerned that our critical incidents may not have been typical, but we do not feel that this is a problem. Each incident included approximately five decision points that were probed, and many of these were routine. In addition, had we looked only at critical incidents, this would have overemphasized the difficult and deliberated decisions, not the prototype matches.

The RPD model of decision making under time stress has potential implications for the selection, training, and support of decision makers. The model suggests that it will be useless to require decision makers to make comparative evaluations of several options. This is a time-consuming process. Instead, decision makers must rely on their experience and ability to quickly recognize the causal dynamics of situations as a way of generating effective options and evaluating them. With regard to selection, there may be individual differences in the way people feel comfortable relying on recognitional matching as opposed to more analytical processes. If these differences could be established and validated, it might help assign individuals to conditions where analytical evaluations are necessary versus those where analytical evaluations are not possible.

With regard to training, it may be valuable to reconsider the worth of teaching people to generate and evaluate a variety of options if they are going to be placed in situations where this will be counterproductive. Instead, it may be more efficient for training programs to be reconceptualized to emphasize the perceptual learning needed to make fine discriminations and the array of experiences needed to develop situational awareness skills and to acquire a repertoire of options. We will need to perform additional research into the dynamics of situational awareness in order to develop more definite guidelines for training programs.

Finally, the RPD model suggests that it would be a mistake to develop decision aids along the lines of only decision analytical theories. In time-pressured situations, people will not be able to perform the operations needed to make comparative judgments. It would be much more valuable to make sure that decision support systems are providing an effective situational awareness. We hypothesize that decision supports that provide options will not be used, or will lead to degraded performance under time pressure, whereas decision supports for situational awareness will improve performance. This hypothesis may interact with cognitive styles and with the experience level of personnel, but it should be considered prior to the development of support systems for personnel who will be required to make time-pressured decisions.

This effort was extremely valuable for the opportunity to learn about decision making under time pressure. We had expected to study how options are chosen from among alternatives, and instead we found that comparative option selection does not often occur. We wanted to study analogical reasoning, and instead we found little evidence for the direct use of analogs. For these reasons, we were forced to develop a new understanding of decision making, as presented in the RPD model, whereby effective options are directly generated and evaluated for adequacy without any comparisons with other options. This model links decision making to pattern matching, perceptual learning, and the formation of prototypes in memory. It asserts that decision behavior can no longer be appreciated in isolation from these other aspects of psychological functioning.

Future research is needed into the knowledge elicitation tools we used and into the postulates of the RPD model—use and nature of prototypes, aspects of situational awareness, the existence of action queues, conscious analysis of one option at a time, and evaluation of options. It is hoped that the elaboration of the RPD model will provide a means for increasing the applicability of decision research for operational problems and requirements.

## Postscript by Gary Klein, November 2010

I appreciate this opportunity to publish the results of the initial firefighter study conducted by Klein, Calderwood, and Clinton-Cirocco that generated the recognition-primed decision (RPD) model. This study was never even submitted for publication because we were not aware of any journal that would consider it. (We did present the results at the 1986 Human Factors Society conference.) Now, 25 years later, it can finally make its debut in this special issue of the *Journal of Cognitive Engineering and Decision Making,* the rationale of which is to showcase cognitive field research studies such as this one.

When the announcement for this special issue on naturalistic decision making was issued, Robert Hoffman saw the opportunity for publishing the firefighter study and transformed the original 1985 report so that it could be submitted. However, some modifications have been necessary because the original report was 97 single-spaced pages. Robert deleted the appendices (65 pages), but the manuscript was still too long. Therefore I have further shortened the manuscript, primarily by deleting some of the examples and eliminating some of the unnecessary tables.

Frankly, I found the writing a bit clunky—after all, this was a final report and had not been polished to meet journal standards. But I have gritted my teeth and left it as is. I have not changed anything from that 1985 report other than deleting some textual description of the material in the tables and adding a very few transition phrases to bridge gaps left by the deletions. At the request of the editor I made some additional changes to improve readability and to bring the manuscript into conformance with current guidelines: breaking up some long paragraphs, combining some short paragraphs, adding some clarifying comments in brackets, using *Participants* instead of *Subjects*, and combining the reference sections from the original article and this postscript. I made only two small substantive changes. I altered the von Clausewitz reference to reflect its original publication date of 1832, and I corrected in the text a minor error in the original report's Table 2 (which has been deleted for this article). The error did not affect the reported result that 78% of the decisions we studied were estimated to have been made in less than 1 min. Additional details of the methods can be found in Klein (1998).

Here are some of the features of the report that I would change if I rewrote it today. First, I would update the decision model. The RPD model described in the report contained the core of our insights but had not yet been elaborated with a few feedback loops. About a decade later we added a situation awareness function (Kaempf, Klein, Thordsen, & Wolf, 1996) as Level 2 of the RPD model and noted that decision makers seemed to contrast alternative accounts of a situation, even though they rarely contrasted alternative options of what to do.

Second, the report emphasized decision-making research and neglected other research strands, so I would expand on the studies of expertise and knowledge engineering that converged in the field of naturalistic decision making. Third, the report noted one incident that was an outlier—a pumping station fire that took a week to contain and defeated the best efforts of the volunteer fire departments that were called in initially. The report attributed the different pattern of results for this incident to difficulties in organizational dynamics, whereas today I would also emphasize the lack of expertise of the firefighters.

Fourth, the report described an initial step in the development of a critical decision method for doing cognitive task analysis; today I would describe the ways in which that method has evolved (e.g., Crandall, Klein, & Hoffman, 2006). Fifth, I would revise the scoring method because in later studies we concentrated on difficult decisions rather than all the decision points we could identify in an incident. Sixth, the report describes a strategy of using imagery to evaluate options—we subsequently came to describe that strategy as mental simulation (Klein & Crandall, 1995). Seventh, I wish we had explained that the RPD model is not simply about intuition but is a blend of intuition (the prototype matches, which today would be described as pattern-matching) and analysis (the mental simulation).

Nevertheless, I think the report holds up fairly well after 25 years, and I am proud of the research we did. We encountered an unexpected finding—time-pressured decision making without generating and comparing multiple options—and we made sense of it. Also, we appreciated the importance of situational awareness for this type of decision making. I am not sure how we arrived at the term "situational awareness"; I believe that only the aviation community was using that term at the time, and I do not think we were in touch with that literature. In addition, our research team, which included two psychologists, a communications specialist, and an anthropologist, struggled with the boundaries of what we called "quasi-naturalistic" research; several years later the naturalistic decision-making movement was started at a September 1989 meeting in Ohio (see Klein, Orasanu, Calderwood, & Zsambok, 1993). Our firefighter project was funded by a 6-month $50,000 Small Business Innovative Research contract. It was an exciting 6 months.

## Acknowledgments

This paper was originally a Report to the Army Research Institute for the Behavioral and Social Sciences, November 8, 1985 (Klein, Calderwood, & Clinton-Cirocco, 1985). It was subsequently issued in 1988 by the Defense Technical Information Center (DTIC) as an Army Research Institute Technical Report. The authors wish to thank the contract monitor, Dennis Leedom, for his insights, helpful suggestions, and support of this project. Thanks also to Christopher P. Brezovic and Jim Malarkey, who helped with interviewing and coding and contributed to the discussions of study design and interpretation. Helen Altman Klein, Don MacGregor, and Robert Holt provided consultation and helpful criticisms of earlier drafts of this report.

The administrative staff of the Klein Associates Division of Applied Research Associates assisted in producing this updated article based on the original report, as did Veronica Sanger. Robert Hoffman graciously prepared it for publication 25 years after its original appearance.

Finally, our sincere appreciation is expressed to the officers in the Ohio and Indianapolis fire departments who gave generously of their time and expertise to make this study possible.

## References

Aumann, R. J. (1964). Subjective programming. In M. W. Shelley & G. L. Bryan (Eds.), *Human judgment and optimality* (pp. 217–242). New York: Wiley.

Bettman, J. (1971). A graph theory approach to comparing consumer information processing models. *Management Science, 18,* 114–129.

Brandt, R. M. (1981). *Studying behavior in natural settings.* New York: University Press of America.

Campbell, D. T. (1960). Blind variation and selective retention in creative thought as in other knowledge processes. *Psychological Review, 67,* 380–400.

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

de Groot, A. D. (1978). *Thought and choice in chess.* New York: Mouton. (Original work published 1946).

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.

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.

Howell, W. C. (1984). *Task influence in the analytic intuitive approach to decision making* (Final Rep. No. ONR Contract N00014-82 C-001 Work Unit NR197-074). Houston, TX: Rice University.

Kaempf, G. L., Klein, G., Thordsen, M. L., & Wolf, S. (1996). Decision making in complex command-and-control environments. *Human Factors, 38,* 220–231.

Klein, G. A. (1998). *Sources of power: How people make decisions.* Cambridge, MA: MIT Press.

Klein, G. A., Calderwood, R., & Clinton-Cirocco, A. (1985). *Rapid decision making on the fire ground* (Tech. Rep. No. TR-85-46-12). Alexandria, VA: U.S. Army Research Institute for the

Behavioral and Social Sciences. Subsequently released in 1988 as DTIC Tech. Rep. No. AD-A199492, available from http://www.dtic.mil.

Klein, G. A., & Crandall, B. W. (1995). The role of mental simulation in naturalistic decision making. In J. Flach, P. Hancock, J. Caird, & K. Vicente (Eds.), *The ecology of human-machine systems* (pp. 324–358). Hillsdale, NJ: Erlbaum.

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

Kleinmuntz, B. (1968). The processing of clinical information by man and machine. In B. Kleinmuntz (Ed.), *Formal representation of human judgment* (pp. 149–186). New York: Wiley.

MacCrimmon, K. R. (1968). Descriptive and normative implications of the decision theory postulates. In K. Borch & J. Mossin (Eds.), *Risk and uncertainty* (pp. 3–32). New York: St. Martin's Press.

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

Rouse, W. B. (1978). Human problem solving performance in a fault diagnosis task. *IEEE Transactions on Systems, Man, and Cybernetics, 4,* 258–271.

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

Tulving, E. (1972). Episodic and semantic memory. In E. Tulving & W. Donaldson (Eds.), *Organization of memory* (pp. 381–403). New York: Academic Press.

Tversky, A. (1972). Elimination by aspects: A theory of choice. *Psychological Review, 79,* 281–299.

von Clausewitz, C. (1976). *On war.* Princeton, NJ: Princeton University Press. (Original work published 1832).

Zakay, D., & Wooler, S. (1984). Time pressure, training, and decision effectiveness. *Ergonomics, 27,* 273–284.

**Gary Klein,** Ph.D., is a senior scientist at MacroCognition LLC. Dr. Klein received his Ph.D. in experimental psychology from the University of Pittsburgh in 1969. He was an assistant professor of psychology at Oakland University from 1970 to 1974 and worked as a research psychologist for the U.S. Air Force from 1974 to 1978. In 1978 he founded his own research and development company, Klein Associates, which was acquired by Applied Research Associates (ARA) in 2005. He joined MacroCognition LLC in 2009.

**Roberta Calderwood,** Ph.D., is a senior analyst with Science Applications International Corporation (SAIC) performing assessments of new military technologies and operational concepts for a variety of Department of Defense agencies. Prior to joining SAIC in 1990, Dr. Calderwood worked at Klein Associates (1985–1990) on research related to naturalistic decision making. She received her Ph.D. in experimental psychology from the University of New Mexico in 1990.

**Anne Clinton-Cirocco,** M.S.W., L.C.S.W., is the division officer for social work at the Naval Health Clinic Great Lakes. Ms. Clinton-Cirocco received her M.S.W. from California State University, Sacramento, in 1992. She was a graduate research associate with Klein Associates (1985–1986) while pursuing graduate work in applied behavioral sciences at Wright State University. Ms. Clinton-Cirocco has been a psychotherapist and medical social worker in the United States and overseas in civilian and military settings (1992–2010).