A LITERATURE REVIEW OF ANALYTICAL AND NATURALISTIC DECISION MAKING

TASK 2

FINAL TECHNICAL REPORT

Prepared by:

Caroline E. Zsambok, Ph.D

Lee Roy Beach, Ph.D.

Gary Klein, Ph.D.

Klein Associates Inc.

582 E. Dayton-Yellow Springs Road

Fairborn, OH 45324-3987

Prepared for:

Naval Command, Control and Ocean Surveillance Center

Research, Development, Test, and Evaluation Division

271 Catalina Boulevard

San Diego, CA 92152-5000

Date Submitted: 31 December 1992

Table of Contents

List of Tables

List of Figure

Introduction

This is a report about the second of three tasks that Klein Associates completed to examine how the Naturalistic Decision Making (NDM) approach can be applied to designing human-computer interfaces (HCI) and decision support systems (DSSs). We performed all three tasks under a project sponsored by the Office of Naval Technology, called Tactical Decision Making Under Stress (TADMUS).

TADMUS was designed to learn how Naval officers handle very difficult decisions under conditions such as time pressure and uncertainty. Prior to TADMUS, the research emphasis had been on high intensity combat conditions. TADMUS was directed at low intensity conflict (LIC), including high degrees of ambiguity about the nature of a threat, and the intent of a track. The use of AEGIS cruisers in the Persian Gulf during the Iran-Iraq war was an example of this. AEGIS cruisers were designed for blue-water operations, yet they were needed in the Gulf, within very narrow confines, and they lacked some important features for self-defense.

TADMUS is a project aimed at understanding how officers make decisions in a LIC environment, in order to help either with better training of teams and individuals, or with the design of better HCIs or DSSs. Klein Associates began work in support of TADMUS in September, 1990. Our intent was to find ways of designing HCIs and DSSs to improve decisionmaking, building on our past work in Naturalistic Decision Making (NDM) (e.g., Klein, 1989). Previous research on classical, generally analytical, decision strategies has not yielded useful insights for developing better systems for this environment. The question driving this effort was whether a naturalistic decision perspective would do any better.

Our work consisted of three tasks. In Task 1, we conducted interviews with AEGIS commanders and Anti-Air Warfare officers, to study the way they make decisions. The results are described in a separate report (Task 1 Technical Report: Kaempf, Wolf, Thordsen, & Klein, 1992). In Task 2 (reported here), we surveyed the field of classical and naturalistic decision strategies, to see if there are useful ideas to be incorporated into TADMUS. The results of this task are described in the present report. The third task was to draw on both of these efforts to generate a decision-centered approach to designing interfaces and system supports. This task, and the storyboards we developed, are described in a separate report (Task 3 Technical Report: Miller, Wolf, Thordsen, & Klein, 1992). Task 4 is an overview report of the work conducted in the first three tasks (Klein, 1992).

This report is organized in three parts, in addition to the Introduction. First, we present a review of the literature on decision strategies for screening and choice. This literature concerns strategies available to decisionmakers when their task involves selecting one course of action (i.e., "option") from several possible ones. In the second part, we differentiate these strategies from naturalistic decision processes that do not involve option selection. We review a number of models of naturalistic decisionmaking and describe several processes used by decisionmakers to diagnose the situation and to develop a course of action. Finally, in the Conclusion, we summarize the perspectives from both of these sections and offer a matrix of decision processes and strategies that are likely in the context of interest to the TADMUS project.

Part 1: Decision Strategies For Screening And Choice

Since its beginning (Bernoulli, 1738; Pascal, 1670), decision theory has provided strategies to guide decisionmaking. Each strategy has its roots in a common logic and yet is unique in that it is crafted to suit a particular set of circumstances, which it captures in its assumptions and in the procedures that it prescribes for its application. The purpose of

Part 1 is to present the most commonly studied strategies so that we can evaluate their appropriateness for guiding the design of decision support systems for U.S. Navy operational personnel.

The common logic underlying the different decision theoretical strategies regards decisionmakers as `rational,' i.e., it is assumed that they will not intentionally select a course of action, an option, that is inferior to some other option. Options are seen to possess attributes that will, to one degree or another, promote accrual to the decisionmaker of various outcomes should that option be selected. The worth of an option to the decisionmaker is a function, usually additive, of the worth of the outcomes that will be promoted by the option's attributes.

Within these general constraints the details vary. For example, the basic model, maximization of expected value (EV) assumes that there is an explicit set of options and that each option in the set has identifiable potential outcomes, each with a specified value (desired or undesired) to the decisionmaker and a known probability of accruing to the decisionmaker if he or she were to choose that option. The strategy consists of summarizing the value of each option as the sum of the values of its potential outcomes, each discounted by (multiplied by) the probability that the outcome would in fact be obtained. This product sum, called the option's expected value, is then compared with the expected values for the other options. The option that has the largest expected value is the one that should be selected, called `maximization of expected value.'

In contrast, the lexicographic strategy (LEX) assumes that each option has attributes that will promote valued outcomes. Unlike EV, however, LEX ignores probabilities and it does not require summarization of what the decisionmaker knows about each option. Rather, it prescribes that one attribute be selected and the option that is best in terms of that attribute be chosen.

Other strategies make other assumptions and prescribe other procedures. However, in addition to the logical underpinnings described above, each of them assumes that the decisionmaker's goal is either to screen out objectionable options or to choose one best option from among a set of options. In addition, they each assume that the decisionmaker has preferences for the outcomes and that these preferences can be measured. And, they each assume at least a modicum of analytic and mathematical ability on the part of the decisionmaker or on the part of an agent who is acting for the decisionmaker (e.g., a decision aiding system).

The Present Goal

It must be emphasized that the goal of what follows is descriptive and not prescriptive. The logic of the strategies will be outlined, the most crucial and distinguishing procedural differences will be listed, and examples will be provided to give the general idea of how the strategy works. This is in contrast to a prescriptive goal, or even a predictive goal. That is, we do not suggest that these descriptions necessarily delineate the conditions under which a given strategy should be used, nor do we suggest that these are the conditions under which decisionmakers normally use it (a point to which we will return at the end of

Part 1). We merely present the strategies commonly found in the decision literature that have been designed to meet particular goals and to fit particular constraints.

Getting Some Terms and Concepts Straight

Decision and decisionmaking are not used with precision in common parlance. This badly obscures the precise

nature of the process by which decisions are defined and options are selected. In order to think and talk more clearly about this process, we consistently will use <u>decision</u> and <u>decisionmaking</u> as umbrella terms. The more precise term, <u>option</u> <u>selection</u>, will be used for decisions that involve specific strategies that consider the <u>pros</u> and <u>cons</u> (the merits and demerits, the strengths and weaknesses, the benefits and costs, the positive and negative consequences) of options either to <u>screen</u> out undesirable options or to make a <u>choice</u> of the best or at least on acceptable option. Thus, one might decide that one is hungry, but this decision does not result from use of a specific strategy that considers either the pros or the cons of alternative courses of action. Hence, the diagnosis of one's alimentary state, while perhaps properly regarded as a decision, would not be properly regarded as an option selection. Once the diagnosis is made, however, one might well engage in option selection in order to remedy one's hunger.

Although the differentiation between decisionmaking in general and option selection in specific may seem a bit artificial, it will be of crucial importance in this and subsequent sections.

<u>A Matrix for Option Selection Strategies</u>

Table 1 contains the array of option selection strategies that have been identified in the decision literature--both those that derive from normative theory and those that are similar in logic and derive from kindred viewpoints. It includes strategies for screening and for choice. The body of the table summarizes the conditions for which each strategy has been designed.

Table 1. A Matrix for Option Selection Strategies

<u>Conditions</u>

The relevance of an option selection strategy depends on the conditions under which the selection is to be made. It is convenient to divide the conditions into four categories: (1) the goal of the option selection, (2) the characteristics of the available information about the option or options, (3) the application requirements for using a particular strategy for the decision problem of interest, (4) the nature of the environment in which option selection is to be made.

(1) <u>The goal of option selection</u>. Option selection arises when neither the environment nor experience prescribe the next step in performance. In the absence of a prescription, the decisionmaker must consider what is to be done. This may involve a goal of:

• screening out unacceptable options

• selection of an acceptable option

• selection of the best option

(2) <u>Characteristics of the information</u>. Option selection requires the decisionmaker to be at least minimally informed about each option's attributes and the implications of these attributes for the future--about the outcomes the attributes will promote. This information may be presented at the time the option's availability becomes known, it may be retrieved from the decisionmaker's store of knowledge about options, it may be obtained

through inquiry, or it may be obtained from all three sources. Information may merely be about whether the option possesses a particular attribute (called `nominal' measurement of the attribute) or it may be about the extent to which the option possesses the attribute (called `ordinal or better' measurement). Moreover, the information may be about whether an attribute is desirable or undesirable (pros and cons), about the uncertainty of the information

(reliability), and about the completeness of the information. This results in the following conditions for information characteristics:

• measurement properties of the information

- possesses attribute or not (nominal)

- extent (ordinal or better)

• information about desirability of attributes

- desirable (pros)

- undesirable (cons)

• reliability of the information

- reliable

- unreliable

• completeness of the information

- complete

- incomplete

(3) Application requirements. Some option selection strategies are specifically designed to identify a single, final option in a single application (although the decisionmaker may elect to gather more information and repeat application if a single option is not identified). Some are specifically designed to be applied iteratively, eliminating options at each step and finally emerging with a set of roughly equivalent options or with a single best option.

Some strategies require that options be evaluated in reference to the other options that are being considered (relative evaluation), while others require evaluation in reference to a set of standards or criteria (absolute evaluation). When relative evaluation is required, some strategies require that relevant information be organized as a summary of each individual option's attributes (within-option evaluation). Others require that relative evaluation be organized by attributes, resulting in comparisons of the various options on each attribute (between-option evaluation).

Some strategies require compensatory evaluation, which means that the net desirability (pros) or undesirability (cons) of the various attributes is the central focus of evaluation (desirable attributes compensate for undesirable attributes), whether within each option or between options. Others require noncompensatory evaluation, focusing solely upon desirable attributes or solely upon undesirable attributes. These considerations lead to the

following list of (not necessarily independent) conditions for strategy application requirements:

• single strategy applied only once

• single strategy applied iteratively

• use of a second strategy (or reapplication of the same strategy) to break ties when first strategy does not produce a definitive selection

• relative evaluation

-within-option evaluation

-between-option evaluation

• absolute evaluation

• compensatory evaluation

• noncompensatory evaluation

(4) Environmental circumstances. Option selection is made in the context of decision problems that differ to some degree from one time to another. If circumstances remain reasonably constant we can speak of a typical environment within which a particular problem (and option selection process) arises. Thus, personnel hiring problems may arise in many different ways, but within an organization there is likely to be a good deal of overlap in circumstances from one occasion to another. It thus would be prudent to use the same option selection strategy on each occasion.

While it is impossible to list exhaustively the circumstances that might prevail in all problem and selection environments, there are three that are particularly salient, pervasive, and discriminating in terms of selection strategy determination. First, there may or may not be a selection-aiding technology available--anything from a pencil and paper to a computer-supported decision program--and there may or may not be any need to use it if it is available. Second, the problem may or may not be structured by the environment. In some environments the problem is clearly structured in that the goal of the decision is clear, the outcomes are obvious, and the options are easily evaluated. In other cases the goals and outcomes may be ambiguous and the options may be so diverse and complex that they are difficult to evaluate. Third, there may or may not be ample time to apply any strategy the decisionmaker determines to be appropriate, or there may be moderate or severe constraints that rule out use of some strategies. These considerations lead to the following list of environmental conditions:

• availability and advisability of a decision aid, with the need for an aid increasing as the number of options, and the number of attributes associated with each, increases

• degree of problem structure, with structure increasing as clarity about the goal of the decision, specificity of the outcomes, and the ease of attribute evaluation all increase

• time available to make the decision, with time requirements for different strategies being indexed as little, moderate, or extensive

Option Selection Strategies

There are many strategies for making option selections, from tossing a coin or reading Tarot cards to using a computer-aided decision program or employment of a professional decision analyst. However, following the lead of Svenson (1979), the strategies to be examined below are limited to those that can be rationalized within the prevailing logic of mainstream decision theory and research. It also must be noted that there are variations on all of these strategies, some consisting of a merger of two or more of the strategies that will be discussed, some consisting of alterations in the assumptions or procedures of the strategies. It would be impossible, and redundant, to treat each variation as a separate strategy. Therefore, only when a variation has received considerable attention in its own

right have we included it in our list (e.g., both Expected Value and Subjective Expected Utility).

In what follows, each strategy is described in terms of the procedure required to apply it and the conditions that constrain applicability. Just as there are variations on strategies that might themselves be regarded as strategies, there are adaptations of strategies to circumvent various constraints. Again, we cannot include all of these adaptations, so only when they have received considerable attention in their own right have we included them. Too, some conditions are common to all of the listed strategies (e.g., clear organizational or

personal values, measurable attributes) and thus do not help in differentiating among them; such conditions are omitted from the descriptions.

The following descriptions are summarized in Table 1.

(1) Maximization of Expected Value (EV): For each option compute the product sum, across outcomes, of the value of each outcome resulting from selecting that option and the probability that the outcome will in fact occur should the option be selected. Select the option with the largest product sum, i.e., the highest expected value (Bernoulli, 1738). Variations on maximizing include minimax, maximin, and other option selection rules.

Conditions: Designed to choose the best option; requires ordinal or better information; considers both desired (pros) and undesired attributes (cons); takes uncertainty about the information into account; requires complete information about the attributes of the options; is a single strategy applied only once; is used for relative, within-option evaluations; is compensatory; usually requires a decision aid--if only pencil and paper; requires a

well-structured problem; and, depending upon the problem, requires extensive time for application.

Example: The medical director of a mining company's health clinic must select a treatment program for a strain of flu that has resulted in considerable loss of productivity in the mines. She turns to the medical literature for information and finds that three rather different treatments exist. The first reduces the approximately 30 day duration of the illness by 10 days, the second by seven days, and the third by only three days. The first costs $100 per patient, the second costs $85 per patient, and the third costs $27 per patient. The physician calculates the number of miners who are likely to become ill and the dollar amount that the company will loose by their inability to work if they were to go untreated. Next, she calculates, for each treatment, the amount that this loss would be reduced were the treatment used; this savings is the value to the company of using each plan. Then she calculates the product of the number of miners who are likely to become ill and the cost, and the savings, for each treatment plan. Finally, she ascertains the success rate for each plan, the probability that a patient treated using the plan actually will recover in the stated time. She arrives at the expected value of each plan by calculating the product of the probability of a timely cure (P) and the value of using the plan (the savings minus the cost of using it) plus the product of the probability of not obtaining a timely cure (1-P) and the additional cost of having the miners away from work for the full duration of the illness (plus the cost of the failed treatment). She chooses the plan that has the highest expected value.

(2) <u>Maximization of Subjective Expected Utility (SEU)</u>: The same strategy as maximization of expected value except that utility is substituted for dollar value and

subjective probability is substituted for actuarial probability of outcome occurrence (Bernoulli, 1738).

Conditions: Same as for EV.

Example: A newly graduated college student has been offered two attractive jobs and must choose between them. He considers each job's potential positive and negative consequences, evaluates each consequence in terms of its utility or disutility to him, and discounts (multiplies) each utility or disutility by his subjective probability that it actually will accrue to him if he were to choose the job that might yield it. Then he calculates the sum of the discounted utilities (positive numbers) and disutilities (negative numbers) for each job and chooses the job for which the sum is greatest.

(3) <u>Addition of Utilities (AU)</u>: Sum the utilities of each option's attributes and choose the option for which the sum is greatest (Svenson, 1979).

Conditions: Designed to choose the best option, requires ordinal or better information; considers both desired and undesired attributes; does not take uncertainty into account; requires complete information; is a single strategy applied only once; is used for relative, within-option evaluation; is compensatory; usually requires a decision aid; requires a well-structured problem; and, depending upon the number of attributes involved, requires moderate to little time for application.

Example: The college student in the previous example ignores his uncertainty about whether the jobs' attributes actually will yield the anticipated consequences (outcomes) and merely sums the utilities and disutilities of those consequences for each job. Then he chooses the job for which the sum is largest.

(4) <u>Addition of Utility Differences (AUD)</u>: For each attribute of interest, compute the difference between the utility for the attribute for one option and the utility for the same attribute for another option. Then sum the

(weighted) differences and choose the option that the sum indicates to have the higher overall relative utility (Tversky, 1969).

Conditions: Designed to choose the best option; requires ordinal or better information; considers both desired and undesired attributes; does not take uncertainty into account; requires complete information; is a single strategy applied only once; is used for relative, between-option evaluation; is compensatory; requires a decision aid; requires a

well-structured problem; and, depending upon the problem, requires moderate time for application.

Example: Using the job-seeking student again, the utility of a specific consequence for Job 2 is subtracted from the utility of the specific consequence of the same kind for Job 1. That is, the utility for the salary offered for Job 2 is subtracted from the utility of the salary offered for Job 1, resulting in either a positive or negative difference. The utility for the amount of travel necessitated by Job 2 is subtracted from that for Job 1, and so on. Then each positive or negative difference is weighted by the importance of the consequence and summed. If the sum is positive, Job 1 is chosen. If the sum is negative, Job 2 is chosen.

(5) Dominance (DOM): Choose the option the utility of which is at least as attractive as that of every other option for all attributes of interest and better than every other option on at least one attribute (Lee, 1971).

Conditions: Designed to choose the best option; requires ordinal or better information; considers both desired and undesired attributes; does not take uncertainty into account; requires complete information; is a single strategy applied only once but may require use of another strategy to break ties; used for relative, between-option evaluation; is noncompensatory; a decision aid is not usually required; requires a well-structured problem; and, depending upon the problem, requires little time for application.

Example: A personnel officer must choose an employee for an award. The criteria are work attendance, productivity, seniority, and peer ratings. She finds that one employee ties for best with at least one other employee on every criterion except attendance and has a superior attendance record. She selects this employee for the award.

(6) Conjunction (CON): Choose the option that reaches some critical level on all attributes of interest; attribute A and on attribute B and attribute C, etc. (Dawes, 1964).

Conditions: Designed to choose an acceptable option or pool of acceptable options (only incidentally does it screen out unacceptable options); requires nominal or ordinal or better information; considers only desirable attributes; does not take uncertainty into account; does not require complete information; is a single strategy applied once or in conjunction with another strategy to break ties; used for absolute, within-option evaluation; is noncompensatory; does not usually require a decision aid; requires that the problem be well structured only in terms of the specific attributes of interest; and, depending upon the problem, requires little time for application.

Example: A man is thinking about buying a new car. He wants one that costs less than $X and that has four-wheel drive and that gets good gas mileage and that has a good repair record. If he finds only one car that meets all of these criteria, he buys it. If more than one meets them, he uses some other strategy to choose the best car from the reduced set of options.

(7) <u>Disjunction (DIS)</u>: Choose the option that reaches some critical level on attribute A <u>or</u> attribute B <u>or</u> attribute C, etc. (Dawes, 1964).

Conditions: Same as for the conjunctive strategy.

Example: A commercial photographer must choose a male model for a jewelry advertisement. Because she can adapt her photograph to capitalize upon the model's unique characteristics, she will accept a candidate who has a particularly handsome face <u>or</u> an athletic body <u>or</u> expressive hands. If only one model meets one of these criteria, he is hired. If more than one meets them, the photographer uses additional criteria or some other strategy to choose the best model from among the reduced set of options.

(8) <u>Lexicographic (LEX)</u>: [Effectively the same as SFS except that it is iterative] Choose the option that is best on the most important attribute. If two or more options are equal on that attribute, move to the next most important attribute, etc., each time eliminating subordinated options until a single option remains (Fishburn, 1974).

Conditions: Designed to choose the best option; requires ordinal or better information; considers desirable attributes; does not take uncertainty into account; does not require complete information; is iterative; is used for relative, between-option evaluation; is noncompensatory; usually does not require a decision aid; does not require a well-structured problem; and requires little time for application.

Example: A customer considers four television sets with the intent of purchasing one. The most important attribute is price, leading to retention of the two lowest priced sets, which cost roughly the same. These two sets are then compared on the second most important attribute, the third most important attribute, and so on until an attribute is found on which one set surpasses the other, whereupon the superior set is selected.

(9) <u>Elimination by Aspects (EBA)</u>: Select an attribute, perhaps the most important attribute, and eliminate any option that fails to meet some preset criterial level for that attribute. Repeat the process using the next attribute, and the next, etc., each time eliminating subordinated options until a single option remains (Tversky, 1972).

Conditions: Designed to screen out inferior options until a single best option remains; requires ordinal or better information; considers desired attributes, does not take uncertainty into account; does not require complete information; is iterative; used for absolute evaluation; is noncompensatory; requires no decision aid; requires only moderately well-structured problems; and, depending upon the problem, moderate or little time for application.

Example: A university department must hire a new assistant professor. The `short list' contains the names of five candidates. The members of the search committee agree that research productivity is the most important criterion, which eliminates one candidate and reduces the list to four who have equally good records. The committee then turns to teaching effectiveness and eliminates two of the four remaining candidates. Finally, the two survivors are compared in terms of stated willingness to teach in the evening school. The one who is willing is hired.

(10) <u>Number of Superior Features (NSF)</u>: For two competing options, note which one is superior on each feature of interest. Choose the option that has the largest number of `superior to' classifications (Svenson, 1979).

Conditions: Designed to choose the best option; requires ordinal or better information; considers only desired features; does not take uncertainty into account; does not require complete information; is a single strategy applied only once with use of some other strategy to break ties; used for relative, between-option evaluation; is noncompensatory; does not require a decision aid; requires little problem structure; and requires little time for application.

Example: A woman is transferred by her firm to a new city. On a rush visit to the city, she has just one day to rent an apartment to which her furniture, which is in transit, can be delivered. She visits an apartment complex and looks at two apartments. She compares the two apartments and rents the one that is better on the greatest number of relevant features.

(11) <u>Single Feature Inferiority (SFI)</u>: From a pair of competing options, eliminate the one with the least standing on any feature of interest irrespective of the other feature (Svenson, 1979).

Conditions: Same as for number of superior features strategy except that it screens out the weaker option.

Example: A manager has two employees who both want a job that has become vacant. He wants to honor union agreements so he goes to the employment records and compares the seniority of the two employees. He eliminates the one with the least seniority without considering any other of the employees' strengths or weaknesses.

(12) <u>Single Feature Superiority (SFS)</u>: [Effectively the same as LEX except that it is not iterative.] Choose the option that is superior on some feature of interest irrespective of the other features (Svenson, 1979).

Conditions: Designed to choose best option in terms of a single feature; requires ordinal or better information; considers only most desired feature; does not take uncertainty into account; does not require complete information; is a single strategy applied only once or with another strategy to break ties; used for relative, between-option evaluation; is noncompensatory; does not require a decision aid; does not require a well-structured problem; and requires little time for application.

Example: The manager in the previous example wants to put a strong person in the vacant job. So, he asks around to find out what his employees think of the two people who want the job. He selects the person who is generally regarded as the strongest, even though this person is not liked and not regarded as highly on other features.

(13) <u>Single Feature Difference (SFD)</u>: Find the feature on which the options differ most and choose the option that is best on this feature irrespective of the other features (Svenson, 1979).

Conditions: Designed to choose the better option; requires ordinal or better information; considers both desired or undesired features; does not take uncertainty into account; requires complete information; is a single strategy applied only once or with another strategy to break ties; is used for relative, between-option evaluation; is noncompensatory; does not require a decision aid; requires a moderately well-structured option selection problem; and requires little time for application.

Example: The owner of a small business must sign a contract with one of two suppliers. She asks other business owners to describe the biggest difference between the two suppliers and then chooses the one that is described as

best on the feature that was mentioned most often.

(14) <u>Satisficing (SAT)</u>: Choose the first option that meets or exceeds the minimal criteria for some set of features (Simon, 1955).

Conditions: Designed to choose an acceptable option; uses nominal or ordinal or better information; considers only desired feature; does not consider uncertainty; does not require complete information; is a single strategy applied only once; uses absolute evaluation; is noncompensatory; does not require a decision aid; does not require a well-structured problem; and requires little time for application.

Example: The manager of a florist shop needs a clerk for the Easter rush. He puts a `help wanted' sign in the window and hires the first applicant who can use the cash register properly, can take phone orders, and can write legibly.

(15) <u>Satisficing-plus (SAT+)</u>: [An unfortunate name because it is quite different from satisficing in the sense of (14).] Evaluate options on criterial features, eliminating all options that do not meet the multiple criteria. Then alter the cut-offs for the feature and repeat the process; repeatedly alter the cut-offs and evaluate the surviving options until a single option remains (Park, 1978).

Conditions: Screens out inferior options until only one remains; uses nominal or ordinal or better information; considers only desired feature; does not consider uncertainty; is iterative, uses absolute evaluation; is noncompensatory; does not require a decision aid; does not require a structured problem; and requires moderate to extensive time for application.

Example: The flower shop manager in the previous example interviews the applicants for his clerk job but he defers the choice until he has a small pool of applicants from which to choose. He first eliminates all applicants who simply cannot work the cash register, cannot take phone orders, or cannot write legibly. Several applicants remain so he then makes the requirements more stringent and eliminates the candidates who are not highly adept with the cash register, not highly skilled at taking phone orders, and who cannot write an almost perfect hand. He gives the job to the sole survivor--if there is only one. If there is more than one survivor, he raises his standards until there is one.

But Do Decisionmakers Use These Strategies?

Use by decisionmakers of each of the strategies listed above, or at least some version of each of them, has been observed in laboratory studies. This means that, at least under the conditions set up in the laboratory, decisionmakers can use each of these strategies. However, this begs two larger questions: (1) Do decisionmakers spontaneously use strategies such as these as part of their `natural' decisionmaking activity? (2) If not, do they come to use them if they are trained to do so or if they are provided with a decision aiding system that uses them?

Observations of on-the-job decisionmaking find that most decisions are aimed at keeping the organization's efforts generally headed toward some long term goal, rather than aiming at maximization of some immediate outcome (Donaldson & Lorsch, 1983; Selznick, 1957); that only single options are considered at a time, rather than choices being made from among numerous options (Mintzberg, 1975); and that decisionmaking involves considerably more muddling through than it does precise evaluation of crisply defined options that have clearly

delineated features (Cohen, March, & Olsen, 1972). Indeed, while spontaneous decision strategies may vaguely resemble those described above, the former are far more diverse and far less specific than the latter. Of course, `resemble,' vaguely or otherwise, is a matter of judgment. It is not at all clear just how much (and in what ways) a spontaneous decision strategy can differ from a normative strategy and legitimately be regarded as resembling it.

Whether training or decision aids that promote use of the strategies in Table 1 actually do much good also is unclear. Lichtenstein, Slovic, and Zink (1969) tried to train experimental subjects to maximize expected value in a gambling task, to no avail. Decisionmaking courses usually provide instruction in the application of at least some of the strategies described above, but there is no documented evidence that the pupils become generally better decisionmakers as a result. Computerized decision aids almost all use some variant on the EV strategy, but Isenberg (1984) observes that such aids seldom are used, at least by managers. He also observes that even when the aids are used, if the decision prescribed by them is at variance with the decisionmaker's intuition, intuition tends to win.

Decision analysis, which is based on the EV strategy, is broadly used, and decisionmakers sometimes hire decision analysts to do the analysis for them. However, scrutiny of this process suggests that it may not be the decision strategy that helps the decisionmaker so much as the thought and effort invested in clarifying and logically dissecting the problem (Beach, 1990). In short, while it is reasonable to think that the work invested in gathering information and analyzing the problem is of value in decisionmaking, it is not at all clear that much is accomplished by training or by providing decision aids aimed at using that information to analyze the decision problem in the manner prescribed by one of the decision strategies in Table 1.

It is interesting to note that almost all of the strategies in Table 1 were formulated prior to 1979 when Svenson published the paper that this chapter uses as its departure point. Search of the behavioral decision literature provides no new strategies, although there are variants and refinements of those in the table. In part this is because recent theoretical developments and research have focused less on the strategies in Table 1 and more on a major reworking of the SEU strategy called Prospect Theory (Kahneman & Tversky, 1979). In part it is because newer theories of decisionmaking reject the normative tradition altogether. The latter regard decisionmaking as a variety of problem solving, thus bringing it more in line with modern cognitive theory and research (e.g., Beach, 1990; Pennington & Hastie, 1988).

Having described the major strategies from normative decision theory, we turn now to a broader view of decisionmaking. The purpose of Part 2 is to describe other types of decisions besides option selection decisions that occur in many naturally-occurring decision events.

Part 2: Naturalistic Decisionmaking

Overview

The purpose of Part 2 is to review decision strategies and processes identified in naturalistic decisionmaking research, since this line of research relates more directly than does analytic decision research to the way experienced people actually make decisions in operational settings (Beach & Mitchell, 1990; M. S. Cohen, in press; Orasanu & Connolly, in press; Rasmussen, 1986). Naturalistic decision research may offer a fruitful perspective for decision support design in the military tasks of interest to the TADMUS project because it focuses on diagnostic decisionmaking (situation assessment), whereas the strategies discussed previously are relevant to decisions about which option to select.

First, we will briefly discuss the field of naturalistic decisionmaking research. Then, we will describe seven models that either were developed specifically to account for naturalistic decisionmaking, or that are compatible with it. Next, because much of the work in naturalistic decisionmaking concerns assessing a situation as it evolves over time, we will discuss two processes that can affect this assessment: belief updating and seeking confirmation. We will conclude Part 2 with a summary outline of the processes and strategies discussed here.

In the final part of this report we will present and discuss a matrix that combines and summarizes decision strategies described in Parts 1 and 2 which are likely in situations of interest to the TADMUS project.

Introduction

Naturalistic decisionmaking is a relatively new term (Klein, Orasanu, Calderwood, & Zsambok, in press). We use it here to refer to the way people actually make decisions in

their every-day lives, such as in their personal life (buying a car), and on the job (what to do about suspicious readings on a power plant's nuclear reactor).

The reason naturalistic decisionmaking is an important perspective for the TADMUS project is that it adds something new to the study of decisionmaking--something that has relevance in this environment (Orasanu & Connolly, in press). In this section, we will discuss the following elements that are typical in naturalistic decisionmaking:

• situation assessment in addition to option selection

• single option construction and modification (versus generating many options for comparison purposes)

• single option evaluation (versus comparing multiple options to themselves or to a standard)

• changing conditions and ambiguous information versus stable conditions and information within the decision event

• shifting goals versus stable goals within the decision event

• time constraints in deciding what to do

• previous experience by the decisionmaker in the decision event

Considering each of these in turn, previous research on decisionmaking has not emphasized situation assessment--it has largely concerned the strategies available to decisionmakers for option selection, as described in Part 1. But, in addition to option selection, in operational settings people frequently must devote much of their decisionmaking to diagnostic decisions. None of the strategies described previously are relevant to diagnostic decisions.

A second focus of naturalistic decision research is to describe a mode of decisionmaking that people use when they do not wish to compare multiple options. People frequently report that they generated (or recalled) a single option or course of action and then modified it to meet the demands of the situation. Or, they report that they rejected an initial course of action and constructed (or recalled) a new one but did not compare these options to each other nor did they systematically compare each option to a standard as a means to pick the preferred one. Thus, they engaged in a non-comparison mode of option adoption, as opposed to selecting one from a choice set. We will refer to this as a single option evaluation strategy. This finding has been reported frequently (Beach & Lipshitz, in press; Donaldson & Lorsch, 1983; Klein, 1989; Mintzberg, 1975; Peters, 1979).

Single option evaluation not only excludes comparing options to each other (called "relative evaluation" in Part 1) but it also excludes comparing an option, feature by feature, to a standard (called "absolute evaluation" in Part 1). In many operational settings people do not decompose options into features and compare those features to a criterion set, as the strategies described in Part 1 require. This is because the criterion set often is not known, and because the salience of features frequently changes as the decision event unfolds. To evaluate options, people often report that they use other processes (discussed below), like mental simulation, to evaluate whether a planned course of action will work.

Thus, models of naturalistic decisionmaking attempt to address how people can make decisions in situations where the conditions are changing over time, where information is ambiguous, and where the plausibility of potential goals and courses of action is shifting over time. Moreover, naturalistic decision research often concerns real-life situations in which there is limited time to act. This removes the possibility of using many of the decision strategies described in previous research, and substitutes the need to quickly size up the situation in order to construct an acceptable course of action. Last, it is the experience base of the decisionmaker that permits this situation assessment, and the construction and modification of a course of action. Naturalistic decision research aims to model how these experienced decisionmakers function in operational settings.

To avoid potential confusion, we should clarify that naturalistic decisionmaking need not contain all the conditions highlighted above. Some kinds of naturalistic decisionmaking *do* involve multiple option comparison and selection by an inexperienced person, and it can occur in a stable setting without time pressure. One example of this type of every-day decision is the process of buying a car. But many other every-day decisions like those made on the job do involve heavy emphasis on situation assessment, single options or non-comparative option selection techniques, changing conditions, time pressure, and expertise. And naturalistic decisionmaking research offers a perspective to understand these latter types of decisions.

To clarify these two different types of decisions, let us consider the decision event of buying a car versus deciding how to handle a malfunctioning nuclear reactor. To decide which car to buy, the decisionmaker, who does not have a great deal of experience with this situation, probably begins by listing the features he or she wants in the car. Then, the person lists those cars that he or she thinks will meet those criteria and selects one of them using any of the decision strategies described in Part 1. For example, the person might begin by screening out all cars whose cost exceeds $18,000 (strategy: single feature inferiority). Then, of those remaining, the person might choose the ones with an acceptable appearance *and* those whose gas mileage exceeds 25 MPG (strategy: conjunction). Finally, he or she might choose from the remaining pool the car whose dealership was most conveniently located for servicing (strategy: single feature superiority).

Now, consider a case in which the (experienced) operators of a nuclear power plant become aware of suspicious readings from the reactor's monitors. They are confronted with an unusual set of symptoms which they must

diagnose to determine what to do. The decision about the nuclear reactor is different from the car decision for several reasons. First, a great deal of the work in this decision event would occur in the diagnosis phase, in addition to the option adoption phase. The decision about what is wrong is instrumental to the decision about what action to take--the diagnosis decision is the means to the option adoption decision.

The reactor decision differs from the car decision in another way: for the car decision, the situation (the constellation of factors that equate with wanting to buy a new car) remains the same over time, and the pool of possible cars (those available for purchase) and the features they possess remains stable. For the reactor decision, the situation is changing. Certain new aspects of the situation will emerge as the reactor continues to malfunction, old ones can vanish, and the values on the remaining aspects--like water pressure--can change over a short period of time. These changes lead to ambiguity in the situation, and will affect the plausibility of a course of action at various points in time. This equates to shifting, unstable goals for which the pool of possible options (and their features) is not stable over time, as it is in the car decision.

Moreover, this unstable pool of possible options is more hypothetical than real. Based on our research in a variety of operational settings, we have found that experts typically do not compare different options at different points in an evolving situation. Rather, they modify an initial course of action to accommodate changes in the situation. So, in our example, the operators would likely construct an option--a plan--that would include certain beginning steps to ease the danger while creating feedback about the situation. The plan would be adjusted and adapted as the situation changed, in an effort to keep the reactor running while addressing the malfunction. The "features" of the option would not all be known in advance, and they would change over time. Thus, the option would not be decomposed into all its features and compared to a standard. This type of option evaluation is not a part of the description of strategies offered in Part 1.

Further, in those few cases where a course of action can no longer be revised--where the operators decide that an initially generated option needs to be rejected--the new course of action they construct will not be compared to the previous one to decide which is better. Notice that this is not a "selection" decision, as defined in Part 1, since selection assumes the presence of multiple options and the weighing of their pros and cons (based on their features) in order to choose one.

To summarize, in the car example,

• the major decision concerns option selection

• there are several options from which to select one

• conditions in the decision event are stable:

- features of the situation are stable

- features of options are known

- options are stable over time

- goals are stable over time

• there is relatively little time pressure to select an option

• the decisionmaker is not practiced with this particular type of decision.

In contrast, in the reactor example,

• the major decisions are about both situation assessment and course of action

• either a single option is considered or if one is rejected, non-comparative evaluation of additional ones occurs

• conditions in the decision event are subject to change:

- features of the situation are changing

- changes in the situation influence the viability of goals and courses of action

- changes in the situation require modification or a constructed course of action

• there is limited time to make the decision

• the decisionmakers are experienced with this operational setting.

The purpose of the next section is to present a review of literature that helps us understand more about the type of decisionmaking discussed in the reactor example. For purposes of convenience, we will refer to this type of decision event as one in which diagnostic decisions predominate. The other type of decision event is one in which option selection decisions predominate. We will draw heavily from research within the Naturalistic Decisionmaking paradigm, but will include relevant considerations from studies about inference making, problem solving, belief updating, and option generation.

Models Of Naturalistic Decisionmaking

Recognition-Primed Decisionmaking

Klein (1989; Klein, Calderwood, & Clinton-Cirocco, 1986) has developed a model of Recognition-Primed Decisionmaking (RPD) that describes how experienced people commonly make decisions in their operational settings. For ease of description, the model differentiates between highly familiar and moderately familiar situations. However, it does not presume that this dichotomy exists as a psychological reality--it assumes that familiarity with situations is a continuum.

The model describes how decisionmakers are able to draw upon their experience to assess situations and to arrive at a course of action. Based on observations from five field studies in different domains such as firefighting and tank platoon maneuvers, Klein et al. (1986) found that commanders were often able to quickly size up the situation, arrive at a course of action to deal with it, and modify the course of action as necessary to accommodate changes in the situation.

When the situation is highly familiar. Klein et al. (1986) found that most often, experienced decisionmakers find themselves in highly familiar situations. For example, a fire chief arrives on the scene of a house fire. He perceives a number of cues (features) from the environment: He sees smoke coming from under the eaves of a pitched roof, a red flame shooting out the attic window, and a yellowish flame forming at an adjacent second-story window. The model postulates that these perceptions cue the chief's memory for other similar situations he has seen or for a prototypical instance which is the amalgamation of many such situations he has seen. The model does not specify the nature of stored information (examples or prototypes or both). Neither does the model specify whether a feature-by-feature matching process or a more holistic (pattern) matching process accompanies identification of the situation. But, the model's descriptive capability remains functionally the same regardless.

In fact, both types of memory storage and perceptual matching processes are likely. As described by Barsalou (1992), results from learning procedures in connectionist networks allow us to speculate about processes used by people. People can use feature matching, or they might be able to use a small number of features, perceived not individually but as a chunk, to identify situations. This speculation is based on the fact that learning procedures essentially allow connectionist nets to extract prototypical information across many exemplars, while simultaneously storing idiosyncratic (feature) information about individual examples. Because these models store so many relations among detectors, partial patterns of information are often able to activate related properties and inhibit irrelevant ones, resulting in rapid and accurate matches of stimulus information to memorial information.

Returning to the description of the RPD model and our firefighting example, remembered situations (or prototypes) are presumed to contain information about additional *critical cues* to look for (wind strength and direction); about feasible *goals* (saving the property is not feasible, but saving the adjacent house is); about *typical actions* (three separate streams and two converging streams of water will be necessary to save the adjacent house); and about *expectancies* (you should be able to bring the red flame under control within five minutes).

This type of recognition-primed decisionmaking is called a simple match, and is depicted on the left side of Figure 1. Notice that the simple match process produces information about a course of action that the chief can implement automatically if he chooses not to review or evaluate it. Notice also that he gains information about what to expect, given this type of situation. If he begins to see violations to these expectations (perhaps their course of action is not having the desired effect; perhaps, independent of their actions, the situation is evolving in unpredicted ways), these violations will cue the chief to evaluate the actions and to reassess the situation. Both of these processes are depicted on the right side of Figure 1, which concerns complex recognition-primed decisionmaking.

One way to reassess the situation is to reconsider previously encountered cues in light of new ones that are becoming available as the situation unfolds. If the situation then seems familiar--if it is understood to be a categorically different type than the original one--then a new course of action associated with this new situation assessment will become available from memory. This is akin to beginning a second simple match process. But, if the situation cannot be recognized as highly familiar--if it remains only moderately familiar--the decisionmaker engages in complex RPD.

When the situation is only moderately familiar. From the perspective of the RPD model, situations are experienced by decisionmakers as moderately familiar if either

(1) initially, many of the environmental cues do not match memorial information about

Figure 1. Recognition-Primed Decision Model

previously experienced and partially similar situations, or (2) having initially experienced the situation as highly familiar, the decisionmaker begins to notice violations to expectations that were based upon the initial size-up. In the first case--when situations are initially experienced as only moderately familiar--one action available to the decisionmaker is to gather more information. This could involve simply waiting for more information to become available over time. Or it could include looking to other sources for facts or ideas that would allow re-interpretation of the situational cues at hand. In either case, this suggests a wait-and-see tactic--one that is designed to prevent foreclosing on a situation assessment prematurely.

In addition to seeking more information, the decisionmaker can try to reassess the situation. This might be needed if the events were difficult to interpret, or if more than one interpretation was possible. One strategy used for reassessment is feature matching--using the features of the situation to retrieve or build a hypothesis, or to contrast different hypotheses about the situation. The feature-matching strategy is an elaboration of the RPD model based on the work of Noble (1989), see below. A second strategy is to use mental simulation to imagine a sequence of events that might have plausibly resulted in the observed state of affairs. Mental simulation can also be used to evaluate alternate hypotheses, to see which makes the most sense. The mental simulation strategy is related to the work of Pennington and Hastie (1988), also covered below.

If the situation can then be recognized as very familiar, the decisionmaker is involved in a simple match procedure as discussed above. But in either case--when the situation is not initially highly familiar or when there are violations to expectations--if the time for action is drawing close, experienced decisionmakers are still able to act. This is because they can modify a course of action that they retrieve from memory about a similar situation or prototype to accommodate the current situation.

Klein et al. (1986) have studied situations that mostly were unfolding and changing over time. They found that decisionmakers usually did not react to the changing situation by diagnosing an initial part of the event as "Type A Situation," then diagnosing a later part as "Type B Situation," and so on. Rather, they usually understood initially that the situation was a particular type, and that as time went on they could expect it to evolve in particular ways. Thus, these decisionmakers were not calling up from memory separate packets of information about the courses of action, expectancies, goals, and cues associated with each change in the situation. So, a more reasonable description of the decisionmakers' mental activities is that they diagnosed the situation, and then modified an initially remembered course of action to fit the changing situation. In studies by Klein and his colleagues, decisionmakers reported that they continuously modified their planned and actual course of action to accommodate changes that were occurring in the situation.

This description of recognition-primed decisionmaking offers an opportunity to further clarify differences between the decision strategies discussed in Part 1 and those described here. For example, it might appear that, contrary to the above analysis, an elimination-by-aspects (EBA) strategy was being used to arrive at a course of action. The argument would be that a plan or course of action is developed, the first few steps are imagined or are actually executed, the situation changes, and so the plan is modified to accommodate those changes. This modified plan could be said to be a new plan (a new option), and that the previous one was eliminated because several of its features did not meet the (changed) requirements within the situation.

While the above paragraph offers a logical argument, it is our judgment that extending the meaning of EBA to include this circumstance misses the point. EBA was defined in Part 1 as the strategy of selecting an attribute of an option or course of action and eliminating any option that fails to meet some preset criterial level for that attribute, and then repeating the process using the next attribute and so on until a single option remains. EBA is meant to describe a process that decisionmakers can use to pare down the pool of options until only one remains.

Yet, in the situations described above concerning RPD, multiple options were not being compared to a standard or to each other. There was no need to pare down. Further, all the features of the option were not specified in advance--they became known as the event unfolded. The decisionmakers did not identify specific features of the option and compare them to a standard. What these decisionmakers did was to alter a single remembered course of action (or prototype) to fit the specific situation in which they were acting. This type of decisionmaking parallels that found by Mintzberg (1975) who concluded that the business managers he studied almost always considered only one option (Beach & Lipshitz, in press).

In Klein's (1989) studies, they found that one strategy the experienced decisionmakers used to make these modifications was to mentally simulate each step in the envisioned course of action to determine if it would work, then mentally modify those steps that needed changing before executing them. As depicted in Figure 1, recognizing that a situation is only moderately familiar cues the decisionmaker to engage in mental simulation, and not to implement the initially activated course of action without this evaluation. Other conditions that can trigger mental simulation are high stake situations and awareness that available resources will not permit a course of action to be carried out in its remembered form, so that the decisionmaker needs to mentally modify the remembered course of action before implementation.

Other researchers have noted the importance of mental simulation. Gettys (1983), in his research on hypothesis generation, describes a "walk-through" process that is akin to mental simulation. Here, the decisionmaker imagines performing an act, observing its effects, perhaps (mentally) taking another action, etc., until the goal is reached. For a review of the literature on mental simulation, see Klein and Crandall (in press).

While it is true that single options were considered in most decision events Klein (1989) and his colleagues studied, there were some non-routine cases in which more than one option was considered. But in these cases, a non-comparative strategy was used. That is, the decisionmakers did not generate a number of options in order to have several from which to select one. Nor did the decisionmakers compare each feature of the option against a standard. Instead, they used mental simulation to see if the option they were considering would work--if not, and if it could not be modified to seem workable, they rejected it and then constructed another one. Thus, even in the cases where multiple options were considered, an EBA-type strategy is not a plausible explanation for the manner in which options were evaluated, since EBA is a strategy for paring down numerous options and since it requires a feature-by-feature analysis.

The problem here is that we don't know to what extent we can alter the definition of a strategy and expect it to be subject to the same boundary conditions and mediating effects as in the original definition. The problem is not one of *sufficiency*. Yes, it is possible to recast the meaning of EBA so that it suffices as a description of the strategy used by experts

we observed. The question is one of *necessity*: Did they necessarily use a decision strategy akin to EBA?

The same line of questioning is possible for all the decision strategies discussed in Part 1, especially if you relax

certain portions of their definitions. For example, if you assume that decisionmakers don't engage in the explicit mathematical formulations required for SEU--if you assume some sort of mental algebra that occurs without consciously stepping through the math (as some decision researchers have suggested), then you can broaden the reach of each one of these strategies. But once you begin these alterations, it is unclear where to stop to still have the same model. We draw these distinctions because we want to be clear about how we understand certain decision events: We are saying that some of them cannot be reasonably explained away as instances of particular choice strategies. This point is important, since so often this is the approach that is taken by decision researchers.

To summarize the RPD model, the decision processes (which are underlined), and accompanying effects are:

1. Situation Identification

a) Feature matching--situation is judged by matching features of the situation against features stored in memory about previous situations or prototypes

b) Holistic matching--situation is judged by matching larger patterns in the situation against examples or prototypes stored in memory.

c) If situation is experienced as highly familiar,

i) information about other cues, goals, expectancies, and actions is activated from memory

ii) remembered actions do not require evaluation unless expectancies are violated

2. Situation Assessment

Initially, many features of environment do not match features stored in memory about previous situations (or prototype)

a) situation is experienced as moderately familiar

b) diagnosis of the situation requires effort--can occur through mental simulation

c) information about other cues, goals, expectancies, and courses of action from similar situations is activated from memory

d) (optionally) seek more information before acting

e) mentally simulate remembered course of action and mentally modify it before implementation

3. Situation Re-assessment

Expectancies are violated as situation evolves

a) situation is experienced as moderately familiar

b) diagnosis of the situation requires effort

c) (optionally) <u>seek more information</u> before acting

d) <u>mentally simulate</u> remembered course of action and mentally <u>modify</u> it to accommodate evolving situation

4. <u>Course of Action (CoA) Adoption & Construction</u>

a) If simple match, adopt cued CoA (see 1-c and 1-d above)

b) If situation is only moderately familiar, or if expectancies are violated, <u>modify</u> cued <u>course of action</u> (see 2-d and 3-d above)

c) <u>Reject</u> a course of action if it cannot be satisfactorily modified

<u>Noble's Cognitive Model for Situation Assessment</u>

Noble (in press; 1989) has developed a cognitive model to describe the situation assessment portion of a decision event. The model was supported through a series of experiments (Noble, Boehm-Davis, & Grosz, 1986). They found that the model was able to capture the expertise which Navy operators use when they resolve report-to-track problems--when they localize and identify ships from a sequence of situation reports.

In Noble's model, each type of previously experienced problem (or decision event) is treated as if it is stored in memory as a separate <u>reference problem</u>. Noble states that memory storage may not actually correspond to a structure consisting of reference problems and their solutions, but he argues that other plausible structures, such as those containing prototypes, would be functionally similar.

Reference problems contain information about context, goals, solution methods, and other information useful for adapting these solution methods to future problems. He identifies this view as similar to the RPD model, in which it is proposed that goals, expectancies, cues, and actions are contained in memorial representations of experienced situations.

According to Noble's model, if all of the properties of the reference problem match those of a new one, the reference problem becomes "strongly activated." (This corresponds to simple match in the RPD model.) With strong activation, the solution contained in that reference problem would be considered very promising.

Weaker activation occurs when properties of the new problem fail to meet the criteria specified in the reference problem. Here, the decisionmaker would assume that the reference problem's solution method either could not be used to solve the new problem, or that it must be modified before it can be used. (This corresponds to complex RPD.)

If a problem can be solved in multiple ways, it will (weakly) activate several different reference problems in memory. The environmental features of each of these reference problems specify properties (and values of those

properties) that the new problem should have in order for that reference problem's solution method to work. The decisionmaker is thus guided to select the reference problem whose environmental features most closely match those of the current problem. Once selected, the problem solution contained in the reference problem is applied to the current situation.

Thus, the operative decision strategy in Noble's model is feature matching. Noble's model is both deeper and narrower than the RPD model--deeper in that it specifies greater detail about how simple RPD (feature matching) takes place, and narrower in that it does not attempt to describe complex processes like modification or evaluation of a course of action.

Image Theory

Image Theory (Beach, 1990) attempts to explain a wide variety of decision behaviors. We will present that portion of Image Theory which has relevance to our interests in this report. Image Theory is compatible with the RPD model.

Briefly, Image Theory holds that a decisionmaker uses features of the context (the stimulus situation) to probe memory. If the stimulus features of the current context are virtually the same as memorial features, the current context is said to be recognized. However, if stimulus features only resemble them, then the memorial information and all that is associated with it constitute an ad hoc definition of the current context, and it is said to be identified.

Relating terms from Image Theory to the RPD model and Noble's model, in Image Theory, "recognition" is like simple RPD and "strong activation," respectively. It occurs with highly similar matches between stimulus and memorial features. "Identification" is like complex RPD and "weak activation," respectively, which occurs with less similar matches.

Recognition and identification permit the current situation to be framed. A frame is that portion of the knowledge store that the decisionmaker brings to bear on a particular context in order to endow that context with meaning. Frames are updated as events unfold. Their information is the backdrop against which further contextual change is evaluated. Like the RPD model and Noble's model, Beach asserts that once a frame is activated, actions associated with it in memory also become activated and available to the decisionmaker. He calls them policies. If the activated course of action (policy) is not useable in the current context, the decisionmaker must adopt a new course of action, or plan. Plans are devised by consulting the advice of others or reviewing one's own repertory of problem solutions.

Plans are evaluated using the compatibility test. The compatibility test is a strategy that eliminates any option (plan) whose unwanted features exceed the decisionmaker's rejection threshold. Once implemented, plans can be evaluated through progress decisions, which assay the fit between the "ideal future" and the "forecasted future," given that the plan is implemented. Progress decisions allow the decisionmaker to modify portions of a plan to accommodate changes that are encountered in an evolving situation. Progress decisions about plans are like decisions to modify a course of action in the RPD model.

These types of decisions allow the decisionmaker to effectively manage the decision event--to shape the event by attempting to preclude unwanted outcomes, and to alter the effect of events by taking preparatory precautions.

For example, in one of the incidents collected during work on Task 1 under this project (Kaempf et al., 1992), an inbound track remained unidentified as it continued to close on the battle group. In the event that this aircraft had hostile intent, the Commanding Officer (CO) wanted to prevent it from firing--to shape the course of the event. So, the CO decided to "light up" the aircraft to let the pilot

know they were prepared to fire on it unless it either turned away from the battle group or identified itself as friendly on Mode 4 IFF (identify friend or foe).

A second incident contains an example of a preparatory decision. Hostile aircraft were circling an AEGIS cruiser. Even though the captain believed this was meant only to harass them, and that the aircraft did not intend to fire on them, he initiated preparatory action. He ordered his crew to man their guns and prepare their weapon systems. Further, he alerted his crew to watch for a turn-out from the circle, since a sudden change in course might indicate that a missile had been fired, and since a turn-out could often be detected prior to detecting a fired missile.

Both of these examples concern decisions taken to modify a single plan. However, in those cases where experienced decisionmakers do have multiple options available in a choice set, Image Theory asserts that they select one by using the profitability test. The profitability test is not a single mechanism--it can include any of the strategies designed to choose the best candidate, as discussed in Part 1.

The term "profitability test" highlights an important issue about higher-order decision strategies. Any time decisionmakers need to use the profitability test, it is because they have multiple options in their choice set. Sometimes decisionmakers are aware of these options at the beginning of a decision event, as in the example of choosing a car from the set that are available. Other times, decisionmakers are not aware of options and specifically decide to create an option set as a strategy for arriving at the preferred one. In either case, if decisionmakers' over-arching strategy--their "meta-strategy"--requires that they be able to review multiple options in order to select one, we will call this the multiple option meta-strategy. Any of the strategies designed to choose the best or an acceptable alternative, or to eliminate unacceptable ones from the choice set (see Part 1), can be used with the multiple option meta-strategy to select the preferred option.

The multiple option meta-strategy can be contrasted to single-option decisionmaking. Single-option decisionmaking can involve either of two processes, as described in Recognition-Primed Decisionmaking: Decisionmakers consider only a single option during the decision event, or they reject an initial option when it proves unacceptable, forcing them to generate another one which they do not compare to the previous one. We hesitate to refer to this as a meta-strategy because usually decisionmakers who use this mode do not select it as a strategy. Rather, they enter the decision event with the intention of assessing the situation and adopting a course of action that is consistent with that assessment.

In sum, the processes described above from Image Theory, with new ones underlined, are:

1. Recognition: Recognize the situation as virtually the same as remembered information. (Same as simple RPD and Noble's feature matching.)

2. Identification: Construct an ad hoc definition of the situation if it cannot be recognized. (Same as situation assessment in complex RPD.)

3. Plan construction: Formulate a planned course of action if remembered course of action (i.e., "policy") is unsuitable. (Same as modifying an option in the RPD model.)

4. Compatibility test: Eliminate any course of action whose features exceed a rejection threshold.

5. Multiple option meta-strategy: Generate or otherwise produce multiple options, then use a comparison strategy (discussed in Part 1) to select one. Note that one type of multiple option strategy is the profitability test, which subsumes strategies designed to choose the best option (as opposed to either choosing an acceptable option or eliminating unacceptable ones.)

a) compare to standard (called "absolute evaluation" in Part 1)

b) compare to other options (called "relative evaluation" in Part 1)

i) within-option evaluation strategies

ii) between-option evaluation strategies

6. Progress decision: Evaluate whether your course of action will allow you to reach your goal. Modify it if not. (This is like mental simulation and modification of a course of action in RPD.)

A Skill/Rule/Knowledge-Based Model of Cognitive Control

Rasmussen (1983; 1986) has analyzed accidents and simulated decisionmaking of experienced operators of complex automated systems like power plants. His focus is somewhat different from authors of the other models described above in that his intent is to model all portions of peoples' tasks, including sensori-motor behaviors that operators perform unconsciously and effortlessly, as well as those that require conscious deliberation. His model is compatible with those described above.

Based on analyses of these accidents and simulations, Rasmussen developed a model of cognitive control that includes three control levels: skill-based, rule-based, and knowledge-based. Familiarity with a situation (and level of expertise) will determine the level of cognitive control that is exercised and, thereby, the nature of the information used to control the activity and the interpretation of observed information (Rasmussen, 1988). This

model reflects his view of "practical decisionmaking" in which situation diagnosis and action are intimately connected.

Skill-based control represents the highly skilled sensori-motor performance controlled by automated patterns of movements. This type of control is exercised by masters, or people with a high level of expertise. Skill-based control is characterized by the ability to subconsciously generate the movement patterns (actions) required for interaction with the familiar environment by means of an internal, dynamic world model. This internal world model is the cognitive representation of sensory input (signs) from the environment. Experts can read the environment in terms of its affordances (Gibson, 1966), and they can respond to these affordances with the smoothness and harmony of expert craftsmen. They automatically read sensory inputs, which include feedback from their previous actions, so they can adjust their current actions without conscious deliberation (Rasmussen,

in press). Skill-based control is like a program that runs without conscious attention, freeing the person to think of other things.

The next level is rule-based control, characterized by consciously controlled actions. A rule is a procedure or subroutine stored in memory that prescribes actions for a particular situation. There is a fuzzy boundary between rule-based and skill-based behavior that depends on the extent to which behavior is executed automatically or with conscious attention.

For example, while engaged in automatic skill-based control, a person may experience ambiguity or deviation in the environment compared to their internal world model of what the environment ought to be, given this familiar situation. Depending on the nature of this "interrupt," the decisionmaker will consciously scan cues in the environment and associate them to rules for coping. People will review only enough cues to enable them to discriminate among plausible actions (given by rules or procedures). Conscious selection of

actions occurs by mentally reviewing previously encountered situations and rehearsing previous successful actions to evaluate if they will work in the current situation.

Conscious decisionmaking can occur for reasons other than interrupts to skill-based behavior. People can choose at any time to "precondition the required dynamic model." That is, they can consciously recall an analogue and mentally review it to rehearse choice points and prepare themselves for what is likely to come in the current situation. Rasmussen sees this kind of mental operation as similar to Recognition-Primed Decisionmaking, providing there is only one analogue that is reviewed.

The next level of control is knowledge-based, which also involves conscious decisionmaking. This level of control is necessary if people's goals change during a decision event, or if they enter an unfamiliar situation where know-how and rules for control are not available. Here, the control must move to a higher conceptual level in which performance is knowledge based and goal controlled. The goal must be explicitly formulated (or reformulated, in the case of changing goals), based on an analysis of the environment and the overall aims of the person. According to this model, in unfamiliar situations, the person develops different plans to reach the goal. These plans are mentally tested in a trial-and-error fashion through "thought experiments," allowing the decisionmaker to alter plans and select the most successful one. Again, the process of mental simulation as described in Recognition-Primed Decisionmaking is similar to this explanation. But, Rasmussen highlights situations in which people create multiple plans and select the best one, whereas the RPD model depicts single option evaluation.

In sum, the decision processes described in Rasmussen's model concern levels of cognitive control over behavior. They are:

1. automatic (unconscious) control: skill-based behavior

2. conscious control: rule-based behavior

a) recognize cues

b) scan for cues

c) associate cues to task

3. conscious control: knowledge-based behavior

a) analyze situation

b) evaluate and choose goal

c) generate, evaluate, and choose plan

Further, he specifies the process of

4. mental simulation for analogue and plan evaluation.

A Story Model of Decisionmaking

Pennington and Hastie (in press) have developed an explanation-based theory of decisionmaking to account for jury decisions. It focuses entirely on situation diagnosis. It is relevant to our interest here because it is applicable to situations in which the main goal is to evaluate evidence that is complex and uncertain, and for which the implications of its constituents are interdependent. Evaluating evidence to decide guilt or innocence is similar to certain types of situation assessment, like identifying an inbound track as friendly or hostile. We suggest that story construction might be more likely to occur when decisionmakers are faced with highly unfamiliar situations. We further suggest that story construction as described here might be similar to the use of mental simulation for situation diagnosis as discussed under Recognition-Primed Decisionmaking.

In explanation-based decisionmaking, decisionmakers receive information from the environment, and then construct a causal model--a story--that explains the information. The story contains inferences that the decisionmaker generates. Subsequent decisions are based on the story they impose on the information, not just on the information they received. That is, the nature of the story determines subsequent decisions.

According to the theory, the story coordinates three types of knowledge:

• facts or information from the current situation

• knowledge about similar situations

• generic expectations about what makes a complete story, such as believing that people do what they do for a reason

Given a set of known facts in an unfolding situation, knowledge about similar situations, and expectations about what is needed to make a complete story, the decisionmaker can know when important information is missing, and where inferences must be made.

Gettys (1983) identifies a "why" type of hypothesis that decisionmakers often generate to understand an

unfamiliar situation, and which we see as likely to be included in a story about an evolving situation. One example (as described by Beach & Lipshitz, in press) is the incident of the Libyan airliner downed by Israeli planes, where the Israeli crews generated a hypothesis that concerned the "why" of the situation.

Briefly, the situation involved a Libyan airliner that had strayed into Israeli-occupied territory. Prior to actually downing the Libyan airliner, the two Israeli F-4 aircraft sent to intercept it had taken many actions intended to escort it to a safe landing at Rafidim Air Base. There were hand signals from the Israeli crew to the Libyan crew, there were tracer bullets fired, and as a last resort they also fired on the Libyan craft's wing tips. At one point in the episode, the pilot appeared to obey--he descended and lowered the plane's landing gear. But then he turned back in the direction he had come from, as if trying to escape.

Of course, the Israeli pilots knew that in civilian aviation the pilot's responsibility is to protect passengers' safety at all cost. Here was a case of conflicting information: the appearance of a commercial aircraft coupled with flight behavior that was evasive and dangerous. The Israeli crew constructed a hypothesis about what would have to be true in order for this situation to have arisen: This pilot had something to hide from them. Otherwise, he would have landed at Rafidim.

In addition, the crew generated a hypothesis about the situation, which is at the core of the story: This is a terrorist flight in disguise as a commercial flight, and the crew is willing to risk great danger to avoid landing at Rafidim Air Base.

Generating a "why" hypothesis is one of the inferences that people make when constructing a story because it helps to establish relations among facts. Pennington and Hastie (1988) note that stories guide the decisionmaker in understanding the importance of pieces of information because of the hierarchical nature of the episode's representation, and because of the causal structure contained within the story. So, for example, in the above incident, the Israeli fighters interpreted the fact that all the window shades were down as an intentional act, not an accidental one. It was consistent with the notion that the Libyan aircraft had something to hide from them that the window shades would have been put down.

Pennington and Hastie (1988) state that if there is more than one coherent story, then great uncertainty results. If there is only one coherent story, then it is accepted and will be instrumental in reaching a decision about the episode. The greater the story's coverage and coherence, the more acceptable it is and the more confidence a decisionmaker will have in it. Coverage concerns the extent to which the story accounts for evidence. Coherence has three components: Consistency concerns the extent to which the story does not contain contradictions; plausibility concerns the extent to which the story is consistent with real or imagined events in the real world; and completeness concerns the extent to which a story has all of its parts.

This description of story construction is compatible with a model of hypothesis generation by Gettys and Fisher (1979). This model assumes that decisionmakers possess a hypothesis-retrieval process, consisting of directed recursive memory search of long-term memory. They assume the search is triggered by a second process--a plausibility estimation process which assesses the current hypothesis. Their findings supported the conclusion that new hypotheses are generated when information renders the currently held one(s) less probable.

The principles of coverage and coherence provide a fuller description of how decisionmakers evaluate their generated hypothesis or story. They allow us to understand how decisionmakers might diagnose a situation that is not familiar to them--one that they cannot recognize in terms of either simple or complex RPD. Their generic

knowledge of what would constitute a plausible story to explain what they were witnessing would drive them to seek particular information. It would also allow them to hold in check a final judgment if the story lacked coherence, or if multiple stories could be constructed to account for the information.

Thus, in terms of our interests for this report, the decision processes that the story model implies concern specific ways that decisionmakers can understand an unfamiliar situation. To summarize, they are:

1. Construct a story

2. Evaluate story for coverage

3. Evaluate story for coherence

a) consistency

b) plausibility

c) completeness

The SHOR Model

The Stimulus-Hypothesis-Option-Response model (Wohl, Entin, Kleinman, & Pattitati, 1984) represents a command and control (C^2) theory applicable to the military environment. Its purpose is to model the C^2 process, which consists of a "coordinated set

of information-gathering and decision-making functions, carried out with the objective of effective force application" (page 262). The authors state that the C^2 system is hard to describe because it is a dynamic process that frequently occurs under conditions of considerable uncertainty.

The authors developed several SHOR models, each concerning a different aspect of the decision process, such as the anatomy of tactical decisionmaking, and the dynamics of the tactical process. Of particular relevance here is the SHOR model of task elements. These elements are:

• S - stimulus (data) processing

• H - hypothesis generation and evaluation

• O - option generation and evaluation

• R - response, or action

There are two similarities between the SHOR model approach and that of the others we have discussed. First, their motivation was to describe how decisionmakers actually function in an operational setting (here, military command and control). Second, the task element model includes a focus on situation assessment (hypothesis

generation and evaluation).

However, there are notable dissimilarities. First, unlike authors of the other models, Wohl, et al. accept that, ideally, decisionmakers should generate an exhaustive set of hypotheses that are mutually exclusive as a means to arrive at the best course of action

(page 290). Second, they assume that options should be evaluated through an analytic process. For example, they assume that the outcome probabilities of options are weighted by the decisionmaker's often subjective utility assessments which provide the cost (subjective expected utility) or expected net gain. The utilities themselves derive directly from the decomposition of the military objectives. The selection phase would consider options for implementation in the order of their expected net gain.

We agree with Noble and his colleagues (1989) that while the SHOR paradigm made an important contribution to describing military decisionmaking by popularizing the importance of situation assessment, this model did not "break with the outcome calculation tradition of decisionmaking... since the `O' and `R' steps entail estimating the consequences of candidate alternatives" (page 2).

Thus, the model does not take us very far in accounting for findings like Lipshitz's (1988), where Israeli military decisionmakers relied more on situation recognition than on evaluating the consequences of alternatives. Nor does it help us to describe decisionmaking that involves single-option evaluation, as discussed under Recognition-Primed Decisionmaking.

Analogical Reasoning

Analogical reasoning is not a model of naturalistic decisionmaking in the same sense as the models we've described above. Those models aim to describe functional phases of decisionmaking (situation diagnosis; selecting, adopting or modifying a course of action). Analogical reasoning as described here models the way people can use known cases or ideas to help them conceptualize either novel or somewhat unfamiliar problems or situations. As discussed in Klein (1987) and summarized by Eysenck and Keane (1990), analogical reasoning involves mapping the conceptual structure of one set of ideas (called a base domain) onto another set of ideas (called a target domain).

For example, Rutherford (in Eysenck & Keane, 1990) is reported to have used the solar system (base domain) as an analogy that helped him develop his early conceptualization of the atom's structure (target domain). Eysenck and Keane (1990) describe how the mapping in this example might occur, based on findings from research about how people use analogues.

1. Aspects of the base and target are matched. The fact that there are objects in the solar system which attract each other is matched to the fact that there are entities in the atom which attract each other.

2. Aspects of the base--generally relations, such as *revolves around*--are transferred to the target domain. Relations about the planets revolving around the sun are transferred to the atom domain to create the new conceptual structure there: electrons revolve around the nucleus.

3. Coherent, integrated pieces of knowledge are transferred before fragmentary pieces (many of which are never

transferred) from the base domain to the target domain. Integrated knowledge that attraction and weight difference cause the planets to revolve

around the sun is transferred before nonintegrated information about the earth having life on it, which may never be transferred. And, a fourth characteristic of analogical mapping, which does not have a representation in the solar system example is that:

4. Sometimes knowledge is transferred because it is viewed as being pragmatically important or goal-relevant in some respect.

In terms of the RPD model, the Cognitive Model of Situation Assessment, Image Theory, and the Cognitive Control Framework, analogical reasoning is one way that decisionmakers can generate an understanding of only moderately familiar or novel situations. It is also one way that decisionmakers can generate new courses of action in such situations. That is, the fourth characteristic of analogical reasoning--transference based on goal-relevant information--is directly applicable to the Cognitive Model of Situation Assessment. Recall that this model assumes that reference problems are stored in memory. Reference problems contain information about context, goals, solution methods, and other information useful for adapting these solution methods to future problems. Provided that appropriate reference problems are retrieved, decisionmakers would have access to information useful for adapting these solution methods to current novel situations.

Additional Processes in Naturalistic Decisionmaking

In addition to the decision processes discussed above as part of the model descriptions, there are two other processes we would like to highlight--belief updating and confirmation seeking. We have selected these processes for discussion because of their prominence in recent studies within the military domain under conditions compatible with naturalistic decisionmaking. Unlike the discussions accompanying each of the models, we will describe in some detail the studies of interest concerning these processes.

Belief Updating

Belief updating concerns the way people change their judgments or beliefs over time, as they become aware of new information. Belief updating is therefore intimately related to situation assessment, a major focus of this section.

For several decades, a variety of research lines including probabilistic inference, social cognition, and causal inference have investigated whether people's judgments (beliefs) are unduly influenced by evidence they encounter early or late in a series or a decision event. Considerable disagreement has been reported in this and related literature regarding the conditions under which order effects can be expected to influence judgment.

We will discuss recent studies designed to test the Hogarth-Einhorn model (1992) of belief updating, which these authors developed to account for conflicting findings in the literature mentioned above. We will focus on naturalistic decisionmaking studies from which we conclude that:

• the order in which information is presented to decisionmakers can affect their final judgment if they update their belief as each piece of information is encountered

• order effects can be eliminated if decisionmakers wait until all information is received before making their judgment

The Hogarth-Einhorn model (1992) predicts a recency effect when evidence is presented sequentially, and when people are asked to judge the likelihood of a given hypothesis after each piece of evidence is presented to them. They label this the Step-by-Step (SbS) response mode. A recency effect occurs because under these conditions, people form their judgment by anchoring on a current position and then adjusting their judgment on the basis of new information. Both the direction (positive versus negative) and strength of each new piece of information affect the position of the anchor. And, since each new piece of information creates a new anchor, recent information is weighted more than prior information, creating a recency effect. If some of the information is positive (confirmatory), and some of it is negative (disconfirmatory), the order in which it is encountered will affect the final judgment that is made.

In contrast, order effects should not be found when people adopt an End-of-Sequence (EoS) response mode. Here, people form their judgments about an hypothesis after all the information has been presented. As discussed by Adelman, Tolcott, and Bresnick (1991), when evidence is presented all at once and a probability estimate about an hypothesis is obtained at that time, the Hogarth-Einhorn model predicts that people will anchor on the piece of information presented first and adjust it on the basis of the <u>aggregate</u> impact of all subsequent information in support of or against the initial anchor. In the EoS mode, if some of the information is positive (confirmatory) and some of it is negative (disconfirmatory) the order in which it is presented is inconsequential, since it is considered in the aggregate. The EoS response mode can be considered as a global type of assessment strategy, such as an averaging strategy.

Adelman et al. (1991) described several studies involving trained personnel performing their substantive task. These studies produced results that support the Hogarth-Einhorn model. But there are other studies with experienced participants that do not support it. For example, in a series of experiments where information was presented sequentially, instead of finding a recency effect, Tolcott, Marvin, and Lehner (1989) found evidence supporting a primacy effect. Specifically, Army intelligence analysts overweighted the initial information, and they weighted subsequent information less the later it was received.

Adelman et al. (1991) also described the results of their experiment that was designed to test the Hogarth-Einhorn model. Using a paper-and-pencil format, they presented air

defense operators with information about an unknown target either sequentially (SbS) or all at once (EoS). Some of the information was positive and some was negative relative to a

friendly (hostile) identity. They systematically varied the order in which these pieces appeared in both the SbS and the EoS conditions.

In the SbS condition, operators were asked to provide a probability estimate of the target's identity as friendly or hostile after each piece of information was encountered and again after the final piece. In the EoS condition, a probability estimate was obtained at the time that all the information was provided. Their results supported the Hogarth-Einhorn model. When information was presented in the SbS mode, the order in which it appeared significantly affected the final mean probability estimates, resulting in a recency effect. In contrast, when in the EoS mode, the order in which the information appeared did not affect mean probability estimates. This suggests

that operators used a more global strategy in forming their estimates.

In a second study, Adelman and Bresnick (1991) studied Tactical Control Officers who were in a simulator of the Patriot air defense system. Adelman and Bresnick were concerned with the generalizability of the previous study since it involved a paper-and-pencil test. However, they were able to replicate the previous results, which were consistent with the Hogarth-Einhorn model. That is, in the SbS response mode, officers made different identification judgments and took different engagement actions depending on the order in which the same information was presented to them. And, they again found no order effects in the EoS response mode.

The authors reported a number of caveats to the interpretation of these results. Among them were evidence for a confirmation bias and large individual differences in susceptibility to order effects.

First, concerning the confirmation bias, post hoc analyses found that for one particular track, information about the track could be interpreted as either hostile or friendly. They found that what caused an officer to interpret it one way or the other was his initial identification of the track. This represents a confirmation bias, which is consistent with Tolcott's results mentioned above, and which is inconsistent with the Hogarth-Einhorn model. Evidence of discounting was also found in Adelman's first study.

Second, in both of Adelman's studies, they found large individual differences in the susceptibility of order effects when in the SbS response mode, based on the participants' experience. These post hoc analyses imply that when people have more experience in dealing with conflicting information in their area of expertise, they are less susceptible to order effects. We will say more about this in the next section about seeking confirmation.

In sum, the decision strategies described above concerning belief updating are:

1. Anchor-and-adjust: anchor on a current belief strength, and adjust strength and direction of belief as each piece of new evidence is encountered.

2. Global: determine strength of belief based on the aggregate of all evidence.

Seeking Confirmation

There has been a great deal of research into the "confirmation bias"--the tendency for people to weigh more heavily information that supports their hypothesis than information that contradicts it. This research spans areas as diverse as learning, reasoning, decisionmaking, and hypothesis testing. Our aim is to sample portions of this literature that are relevant to decisionmaking in operational settings.

There is an obvious relation between research on belief updating and seeking confirmation. If people form a belief based on early information and also either seek information to confirm that belief or discount information that contradicts the belief, then this primacy effect is strengthened by a confirmation bias.

We will discuss research about the confirmation bias from which we conclude:

• while a confirmation bias is evident in some studies, it can be reduced through

- training

- changes in information content

- allowing people to actively seek information, instead of passively receiving it

• some studies show that experts do seek disconfirmatory evidence

• a confirmatory strategy can uncover disconfirmatory evidence

One of the most relevant lines of research is that of Tolcott and his colleagues. As described by Tolcott (1991; Tolcott, Marvin, & Lehner, 1989), intelligence analysts differed in their early judgments of where the enemy was going to attack. But, regardless of their first estimate, their confidence in this estimate rose as subsequent information was encountered, even though they all received the same information. Thus, the analysts regarded the new information as confirming their early judgments. Tolcott suggested that it was as if they had "created a model or schema of the enemy's plan, and distorted their assessment of new information to fit their models." In subsequent experiments, Tolcott and his colleagues found that they could reduce the confirmation bias by briefly orienting their participants on biases that can occur in judgment, and by providing them with displays that made explicit the uncertainties about enemy unit locations.

Moreover, Tolcott, Marvin, and Bresnick (1989) found that if subjects could actively select the information they wanted, rather than passively receiving it, they were more likely to pay attention to disconfirming evidence. This manipulation is not a minor one, since in

operational settings the ability to actively seek specific types of information, or to select information from an available set, is common among experienced personnel.

Similarly, Serfaty and Michel (1990) found in their interviews with tactical commanders at various levels of expertise that "while novices mostly seek information to confirm their beliefs about the decision situation (confirmation bias), experts mostly seek information to disconfirm theirs. This difference may be supported by the fact that military commanders know that, in a hostile environment, things rarely go according to plan. Their awareness of an intelligent enemy induces them to look for evidence of deceptive operations, and to prepare for these contingencies."

Recently, some researchers have begun to question whether the confirmation bias necessarily leads to poor decisionmaking. Klayman and Ha (1987) state that many phenomena labeled "confirmation bias" are better understood in terms of a general positive test strategy, which can be an effective decision strategy in certain circumstances. A positive strategy involves testing cases that are expected or known to have the property of interest rather than testing those that are expected or known to lack the property.

Klayman and Ha (1987) argue that under some very common conditions--like predicting a minority phenomenon--you are more likely to receive falsification using a positive test strategy than using a disconfirmatory test. When you are investigating a relatively rare phenomenon, the base rate of the target [p(t)] is low and the set of instances that fit some rule other than the hypothesized rule H is high. Finding a t in H is equivalent to obtaining falsification with a disconfirmatory test. An example would be looking for AIDS victims among people believed

not at risk for AIDS as a way to search for falsification of the hypothesized risk factors. These same conditions also mean that the probability of healthy people [non-targets, or $p(t)$] is high, and the population of people with hypothesized risk factors (H) is small. Thus, finding a t in H is equivalent to obtaining falsification with a confirmatory, or positive test strategy. Here, you would be examining people with the hypothesized risk factors. If you have a fairly good hypothesis, $p(t/H)$ is appreciably lower than $p(t)$, but you are still likely to find healthy people in the hypothesized risk group, and these cases are informative.

The main conclusion from this analysis (p. 217) is that it is important to distinguish between two possible senses of "seeking disconfirmation:" (a) testing cases your hypothesis predicts to be nontargets, and (b) testing cases that are most likely to falsify the hypothesis. It is the latter that is generally prescribed as optimal, yet "in situations where the target phenomenon is relatively rare you are probably better off testing where you do expect the phenomenon to occur rather than the opposite. This situation characterizes many real-world problems" (p.225).

In relation to belief updating, the question is whether this conclusion is warranted, given a probabilistic environment where prediction cannot be expected to be error-free, and where there is uncertainty about what constitutes a correct hypothesis. Klayman and Ha (1987) performed a probabilistic analysis similar to its deterministic counterpart in the above example. They conclude that it would be a rational policy to conduct a positive test--one that would stand a chance of providing a large change in confidence that the belief is correct.

Their objective is not to discredit the value of seeking disconfirmatory evidence. But their analyses show that there can also be value in seeking confirmatory evidence. The expert/novice studies described above found that experts do seek disconfirmatory evidence. The unanswered question with regard to the contexts of interest in this report is whether experts should be encouraged to use both a positive test strategy and a negative one, given that time constraints permit both.

In sum, the decision processes described above are:

1. Seek confirmatory evidence when evaluating an hypothesis about a rare phenomenon (or situation assessment) to uncover disconfirming evidence.

2. Seek disconfirmatory evidence when evaluating an hypothesis about a common phenomenon (or situation assessment).

Summary Outline

Combining all the decision processes, strategies and mechanisms described in Part 2 yields the following summary, which is organized around 1) situation identification,

2) situation diagnosis, and 3) adopting a course of action.

• Identifying a Situation

Feature matching. Judging the situation based on matching features of the situation with features of activated examples or prototypes

Holistic matching. Judging the situation based on matching larger patterns in the situation with activated examples or prototypes

• Diagnosing a Situation

Reassessing. Noting violations to expectancies generated through the identification process and modifying the situation assessment as the event unfolds. Or, mentally simulating

the situation to assess the plausibility of its presumed sequence of events and modifying the assessment accordingly.

Belief updating. Generating an hypothesis about the situation and modifying it

1. Step-by-Step--as each piece of information is encountered

2. Globally--with an evaluation based on aggregated information

Information gathering. Actively seeking more information about the situation prior to finalizing a diagnosis.

1. Passive search--watch more of the event unfold

2. Confirmatory search--seek information that would confirm a situation assessment

3. Disconfirmatory search--seek information that could disconfirm a situation assessment

Story generation. Generating inferences about the situation that lead to a meaningful whole in terms of coverage and coherence.

Analogical reasoning. For novel target situation, match some features to those of an analogue and transfer relevant (functional) information to the target to create a new conceptual understanding of it.

• Adopting a Course of Action

Comparative meta-strategy. Generating multiple courses of action whose features are to be compared to a standard or to each other in order to select one. Then, selecting one, using any of the option selection strategies described in Part 1.

Single option evaluation (i.e., non-comparative strategy). Generating and evaluating a single course of action for its sufficiency.

1. Passive generation--through activation of an example or prototype in memory during situation identification

2. Active generation

a) <u>Mental simulation</u>--mentally simulate each step in an activated course of action

i) <u>modifying</u> any portions as necessary to accommodate the situation

ii) <u>rejecting</u> the course of action if it cannot be modified to accommodate the situation

b) <u>Analogical reasoning</u>--transfer to the target course of action

functional relations about the analogue course of action

3. <u>Compatibility test</u>. Eliminating any course of action whose features exceed a rejection threshold

Conclusions

<u>Summary Matrix</u>

The purpose of Part 2 of this report was to describe various researcher's models and theories about the type of naturalistic decisionmaking in which situation assessment is of major importance--where diagnostic decisions predominate. We contrasted this type of decision making to cases where the decisionmaker's effort is not directed at situation diagnosis, and concerns only the process of selecting one option from many, as discussed in Part 1.

Table 2 presents a matrix that is intended to simplify the material presented earlier. It combines and summarizes concepts from both Parts 1 and 2 of this report. For example, the table combines situation identification with situation diagnosis, since this distinction made earlier was necessary only for a finer-grained understanding of the models described in

Part 2 of this report.

The table classifies strategies and processes according to boundary conditions related to their use. That is, the boundary conditions concern the decisionmaker's intentions: to make a situation diagnosis decision versus a course of action decision and, if the latter, whether the decisionmaker intends to review multiple options versus a single option as the method for selecting a course of action. Note that these boundary conditions are at a more general level than those suggested in Table 4 of the Task 4 report prepared under this project (Klein, 1992). The boundary conditions listed in Table 2 reflect an analysis of the decision research literature and are constrained by those findings, while those in Table 4 of the

Task 4 report reflect new hypotheses.

Table 2 lists only five of the option selection strategies described in Part 1 of this report. Again, in an effort to shorten and simplify, we eliminated strategies that are unlikely in situations of interest to the TADMUS project. These unlikely strategies are ones for which the decisionmaker requires a decision aid, extensive time, or the desire to select the best option.

Note that of the situation diagnosis strategies, mental simulation and analogical reasoning can require more time than feature matching. However, they can be accomplished even in limited time and so are included in the

matrix.

Summary Discussion

We have seen that, historically, the most commonly studied decision strategies concern situations where the decisionmaker is faced with several options from which to select a single one. Researchers have approached this problem from the perspective that the task of the decisionmaker is to identify relevant features of options, and then to compare options, based on these features, either to one another or to a standard as the way to arrive at a single Table 2. Matrix of Decision Processes and Strategies of Interest to the TADMUS Project

	Used For		
Process or Strategy	Situation Diagnosis Decisions	Course of Action Decisions	
		Multiple Options Comparative Selection	Single Options Non- Comparative Evaluation
Feature Matching	X		
Holistic Matching	X		
Seeking More Information	X		

Story Building	X	X	
Step-by-Step Belief Updating	X		
Global Belief Updating	X		
Mental Simulation	X		X
Analogical Reasoning	X		X
Progress Decisions			X
Compatibility Test (option rejection test)			X
Comparative Option selection, especially			
Conjunction		X	
Disjunction		X	

Single Feature Inferiority		X	
Satisficing		X	
Elimination-by-Aspects		X	

option. Part 1 identified 15 strategies commonly discussed in the literature that decisionmakers can use to either screen out unacceptable options or to choose one.

While these option selection strategies can accompany naturalistic decisionmaking (e.g., buying a car), they are limited to situations in which the features of options and the options themselves remain stable during the course of the decision event. Further, some of the strategies require moderate to extensive time to carry out, some are designed to choose the best option instead of an acceptable one, some require a decision aid, and some require more than nominal-level scale information about features. These qualifiers render most of the strategies unsuitable for use in situations where there is time pressure, where goals are changing, and where stakes are high--situations of interest in this report.

The option selection strategies that seem feasible in these contexts include conjunction, disjunction, single feature inferiority, satisficing, and possibly elimination by aspects. The first four meet all of the qualifications just described--EBA meets all but the requirement of better than nominal-level scale information. If we relax that qualification, then EBA can be included as one of five option selection strategies we might expect to find in contexts of interest in this report. On the one hand, as noted in Part 2, this is risky business since we don't know to what extent we can alter the definition of a strategy and expect it to be subject to the same boundary conditions and mediating effects as in the original definition. On the other hand, in this concluding section, we do not want to exclude from consideration those strategies for option selection that could be possible in contexts of interest here.

In addition to option selection decisions, in naturalistic decisionmaking there are other types of decisions that frequently are necessary, such as decisions concerning the diagnosis of the situation, or decisions about

evaluating and modifying a single option instead of comparing it to others. Unfortunately, as many reviewers of the decision literature have concluded, not a great deal is known about how people in operational settings construct hypotheses about situation assessment or how they generate, evaluate, or modify options (Abelson & Levi, 1985; Wohl et al., 1984). Further, one of the most prominent lines of research into option generation, conducted by Gettys and his colleagues (Gettys, 1983; Gettys & Fisher, 1979; Gettys, Pliske, Manning, & Casey, 1987), is not too helpful in the contexts of interest in this report. This is because his interests concerned people's ability to generate an *exhaustive* set of hypotheses that can account for the data (which requires considerable time) rather than generating a plausible one that is satisfactory.

Another candidate research line, that of problem-solving research, is also not very applicable here. The work prior to the 1980s concerned well-structured, well-defined puzzle-like problems, which are not like those of interest here. As Cohen and his colleagues (1992) observed, this early work focused on general-purpose strategies such as breaking down a complex problem into simpler components or working backwards from goals to means. These methods have proven insufficient for problem solving in operational settings, especially those that have many possible solutions (the search space is large) and where there is limited time to find a solution. More recent work concerns expert-novice differences in operational settings like computer programming and medical diagnosis. But these studies, like those of Gettys and his colleagues, do not assume time constraints and changing situations, and they emphasize generating a large set of hypotheses that can account for the data from which to select the best one.

Yet, some recent literature is available that addresses situation assessment and option generation and modification in contexts similar to those of interest in this report. In Part 2, we reviewed several models of naturalistic decisionmaking that are compatible with time-pressured contexts in which the situation is changing over time. We described processes that decisionmakers use during situation assessment and during option generation and modification, and we summarized them at the end of that part.

Finally, we introduced this conclusion section with a matrix of strategies that could reasonably be assumed to be candidates for use in situations of interest to the TADMUS project--situations where there is limited time for diagnosis decisions and action decisions. See the Task 1 Report prepared under this project, for a description of what Kaempf et al. (1992) discovered about the decision strategies used by AEGIS Commanding Officers, Tactical Action Officers, and Anti-Air Warfare Coordinators in time-pressured incidents at sea.

References

Abelson, R. P., & Levi, A. (1985). Decision making and decision theory. In G. Lindzey & E. Aronson (Eds.), Handbook of Social Psychology 3rd Ed., 1, 231-309. NY: Random House.

Adelman, L., & Bresnick, T. (1991). Examining the effect of information sequence on patriot air defense officers' judgments. Organizational Behavior and Human Decision Processes.

Adelman, L., Tolcott, M. A., & Bresnick, T. A. (1991). Examining the effect of information order on expert judgment. Organizational Behavior and Human Decision Process.

Barsalou, L. W. (1992). Cognitive psychology: An overview of cognitive scientists. Hillsdale, NJ: Lawrence Erlbaum Associates.

Beach, L. R. (1990). Image theory: Decision making in personal and organizational contexts. London: Wiley.

Beach, L. R., & Lipshitz, R. (in press). Why classical decision theory is an inappropriate standard for evaluating and aiding most human decision making. In G. A. Klein, J. Orasanu, R. Calderwood, and C. E. Zsambok (Eds.), Decision making in action: Models and methods. Norwood, NJ: Ablex Publishing Corporation.

Beach, L. R., & Mitchell, T. R. (1990). Image theory: A behavioral theory of decisions in organizations. In B. M. Staw and L. L. Cummings (Eds.), Research in organizational behavior, 12. Greenwich, CT: JAI Press.

Bernoulli, D. (1738). Specimen theoriae novae de mensura sortis. Comentarii academiae scientiarum impiales petropolitanae, 5, 175-192. Trans by L. Sommer (1954). Econometrica, 22, 23-36.

Cohen, M. D., March, J. G., & Olsen, J. P. (1972). A garbage can model of organizational choice. Administrative Science Quarterly, 17, 1-25.

Cohen, M. S. (in press). The naturalistic basis of decision biases. In G. A. Klein, J. Orasanu, R. Calderwood, & C. E. Zsambok (Eds.), Decision making in action: Models and methods. Norwood, NJ: Ablex Publishing Corporation.

Cohen, M. S., Adelman, L., Tolcott, M., Bresnick, T., & Marvin, F. (1992). Recognition and metacognition in commanders situation understanding. CTI Technical report (1992). Arlington, VA.

Dawes, R. (1964). Social selection based on multidimensional criteria. Journal of Abnormal and Social Psychology, 68, 104-109.

Donaldson, G., & Lorsch, J. W. (1983). Decision making at the top. New York: Basic Books.

Eysenck, M. W., & Keane, M. T. (1990). Cognitive psychology: A student's handbook. Hillsdale, NJ: Lawrence Erlbaum Associates.

Fishburn, P. (1974). Lexicographic order, utilities and decision rules: A survey. Management Science, 20, 1442-1471.

Gettys, C. F. (1983). Research and theory on predecision processes (TR 11-30-83). Norman, OK: University of Oklahoma, Decision Processes Laboratory.

Gettys, C. F., & Fisher, S. D. (1979). Hypothesis plausibility and hypothesis generation. Organizational behavior and human performance, 24, 93-110.

Gettys, C. F., Pliske, R. M., Manning, C., & Casey, J. T. (1987). An evaluation of human act generation performance. Organizational Behavior and Human Decision Processes, 39, 23-51.

Gibson, J. J. (1966). The senses considered as perceptual systems. Boston, MA: Houghton-Mifflin.

Hogarth, R. M., & Einhorn, H. J. (1992). Order effects in belief updating: The belief- adjustment model. Cognitive Psychology, 24, 1-55.

Isenberg, D. J. (1984, November-December). How senior managers think. Harvard Business Review, 81-90.

Kahneman, D., & Tversky, A. (1979). Prospect theory: An analysis of decision under risk. Econometrica, 47, 263-291.

Kaempf, K. L., Wolf, S. P., Thordsen, M. L., & Klein, G. (1992). Decisionmaking in the AEGIS Combat Information Center. Fairborn, OH: Klein Associates Inc. Prepared under contract N66001-90-C-6023 for the Naval Command, Control and Surveillance Center, San Diego, CA.

Klayman, J., & Ha, Y-W. (1987). Confirmation, disconfirmation, and information in hypothesis testing. Psychological Review, 94(2), 211-228.

Klein, G. A. (1987). Applications of analogical reasoning. Metaphor and Symbolic Activity, 2, 201-218.

Klein, G. A. (1989). Recognition-primed decisions. In W. B. Rouse (Ed.), Advances in Man-Machine System Research, 5, 47-92. Greenwich, CT: JAI Press, Inc.

Klein, G. A. (1992). Decisionmaking in complex military environments. Fairborn, OH: Klein Associates Inc. Prepared under contract N66001-90-C-6023 for the Naval Command, Control and Surveillance Center, San Diego, CA.

Klein, G. A., & Crandall, B. W. (In press). The role of mental simulation in naturalistic decision making. In J. Flach, P. Hancock, J. Caird, and K. Vicente (Eds.), The ecology of human-machine systems. Hillsdale, NJ: Lawrence Erlbaum Associates.

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

Klein, G. A., Calderwood, R., & Clinton-Cirocco, A. (1986). Rapid decision making on the fire ground, Proceedings of the 30th Annual Human Factors Society, 1, 576-580. Dayton, OH: Human Factors Society.

Lee, W. (1971). Decision theory and human behavior. New York: Wiley.

Lichtenstein, S., Slovic, P., & Zink, D. (1969). Effect of instruction in expected value on optimality of gambling decisions. Journal of Experimental Psychology, 79, 236-240.

Lipshitz, R. (1988). Making sense of decision making: The implausibility of real work decisions as consequential choice. Boston, MA: Boston University, Center for Applied Behavior Science.

Miller, T. E., Wolf, S. P., Thordsen, M. L., & Klein, G. (1992). A decision-centered approach to storyboarding anti-air warfare interfaces. Fairborn, OH: Klein Associates Inc. Prepared under contract N66001-90-C-6023 for the Naval Command, Control and Surveillance Center, San Diego, CA.

Mintzberg, H. (1975, July-August). The manager's job: Folklore and fact. Harvard Business Review, 49-61.

Noble, D. (in press). A model to support development of situation assessment aids. In G. A. Klein, J. Orasanu, R. Calderwood, and C. E. Zsambok (Eds.), Decision making in action: Models and methods. Norwood, NJ: Ablex Publishing Corporation.

Noble, D. (January, 1989). Application of a theory of cognition to situation assessment. Technical report for the Office of Naval Research, Contract N00014-84-0484.

Noble, D., Boehm-Davis, D., & Grosz, C. G. (1986). A schema-based model of information processing for situation assessment. Vienna, VA: Engineering Research Associates. (NTIS No., ADA163150).

Noble, D., Truelove, J., Grosz, C. G., & Boehm-Davis, D. (1989). A theory of information presentation for distributed decision making. Vienna, VA: Engineering Research Associates. (NTIS No. ADA216219).

Orasanu, J., & Connolly, T. (in press). The reinvention of decision making. In G. Klein, J. Orasanu, R. Calderwood, and C. E. Zsambok (Eds.), Decision making in action: Models and methods. Norwood, NJ: Ablex Publishing.

Park, C. W. (1978). A seven-point scale and a decision maker's simplifying choice strategy: An operationalized satisficing-plus model. Organizational Behavior and Human Performance, 21, 252-271.

Pascal, B. (1670/1965). Pensees: Thoughts on religion and other subjects. New York: Washington Square Press.

Pennington, N., & Hastie, R. (1988). Explanation-based decision making: Effects of memory structure on judgment. Journal of Experimental Psychology: Learning, Memory and Cognition, 14, 521-533.

Pennington, N., & Hastie, R. (in press). Methods of naturalistic decision research. Norwood, NJ: Ablex Publishing Corporation.

Peters, T. J. (1979, November/December). Leadership: Sad facts and silver linings. Harvard Business Review, 164-172.

Rasmussen, J. (1983). Skill, rules and knowledge: Signals, signs, and symbols, and other distinctions in human performance models. IEEE Transactions on Systems, Man and Cybernetics, SMC-13(3), 257-266.

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

Rasmussen, J. (1988). A cognitive engineering approach to the modeling of decision making and its organization in: Process control, emergency management, CAD/CAM, office systems, and library systems. In W. B. Rouse (Ed.), Advances in man-machine systems research, 4. Greenwich, CT: JAI Press, Inc.

Rasmussen, J. (in press). Deciding and doing: Decision making in natural contexts. In G. Klein, J. Orasanu, R.

Calderwood, and C. E. Zsambok (Eds.), <u>Decision making in</u> <u>action: Models and methods</u>. Norwood, NJ: Ablex Publishing.

Selznik, P. (1957). <u>Leadership in administration: A sociological interpretation</u>. Evanston, IL: Row, Peterson

Serfaty, D., & Michel, R. (1990). Toward a theory of tactical decision making expertise. <u>Symposium on Command & Control Research</u>. McLean, VA: SAIC.

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

Svenson, O. (1979). Process descriptions in decision making. <u>Organizational Behavior and</u> <u>Human Performance, 23</u>, 86-112.

Tolcott, M. A. (June 1991). <u>Understanding and aiding military decisions</u>. Paper presented at the 27th International Aplied Military Psychology Symposium, Stockholm, Sweden.

Tolcott, M. A., Marvin, F. F., & Lehner, P. E. (May/June 1989). Expert decision making in evolving situations. <u>IEEE Transactions on Systems, Man and Cybernetics, 19</u>(3).

Tversky, A. (1969). The intransitivity of preferences. <u>Psychological Review, 76</u>, 31-48.

Tversky, A. (1972). Elimination by aspects: A theory of choice. <u>Psychological Review, 79</u>, 218-299.

Wohl, J. G., Entin, E. E., Kleinman, D. L., & Pattipati, K. (1984). Human decision processes in military command and control. In W. B. Rouse (Ed.), <u>Advances in man-</u> <u>machine systems research, 1</u>. Greenwich, CT: JAI Press, Inc.

ACRONYM LIST

AAW Anti-Air Warfare

AU Addition of Utilities

AUD Addition of Utility Differences

CO Commanding Officer

CoA Course of Action

CON Conjunction

DIS Disjunction

DSS Decision Support System

DOM Dominance

EBA Elimination-by-Aspects

EoS End-of-Sequence

EV Expected Value

HCI Human-Computer Interface

IFF Identify Friend or Foe

LEX Lexiocgraphic

LIC Low Intensity Conflict

NDM Naturalistic Decision Making

NSF Number of Superior Features

RPD Recognition-Primed Decisionmaking

SAT Satisficing

SAT+ Satisficing-plus

SbS Step-by-Step

SEU Subjective Expected Utility

SFD Single Feature Difference

SFI Single Feature Inferiority

SFS Single Feature Difference

SHOR Stimulus-Hypothesis-Option-Response

TADMUS Tactical Decision Making Under Stress