ONR Europe Reports

92-4-W

The 27th International Applied Military Psychology Symposium:

A Focus on Decision Making Research

Stanley C. Collyer (with Invited Papers)

4 August 1992

Office of Naval Research European Office

CONTENTS

The 27th International Applied Military Psychology Symposium:
A Focus on Decision Making Research

by S.C. Collyer

INVITED PAPERS

The 27th International Applied Military Psychology Symposium: A Focus on Decision Making Research

by Stanley C. Collyer, Technology Area Manager for Training and Personnel Systems, Office of Naval Technology, Arlington, VA.

INTRODUCTION

The first International Applied Military Psychology Symposium (IAMPS) was held in 1963. Since then it has evolved into an important forum for information exchange among European and North American military psychologists. Attendees are primarily uniformed and civilian psychologists employed by their respective defense establishments. A wide array of topics are typically covered. These include personnel selection and classification, manpower planning, leadership, stress, training, simulation, and human factors. Over the years IAMPS has been hosted by many countries. It has grown in size, although attendance is restricted to provide an atmosphere that encourages discussion. The Office of Naval Research European Office (ONR Europe) continues its role as sponsor and coordinator of IAMPS.

The 27th IAMPS was hosted by Sweden's National Defense Research Establishment and was held in Stockholm, Sweden on June 10-14, 1991. The meeting was attended by more than 50 people from 14 nations.

The National Defense Research Establishment (commonly abbreviated as FOA, from the Swedish *Försvarets Forskningsanstalt*), is Sweden's principal military and civil defense research organization. In addition to psychology, FOA conducts research and development in the areas of physics, chemistry, medicine, and information technology. Behavioral and social science research is conducted in the Department of Human Studies, in such areas as cognition, human factors, human-computer interaction and training.

The keynote speech was delivered by Brigadier General Tode from the Defense Staff of the Swedish Armed Forces. General Tode noted that despite the monumental political changes taking place, aggression is still a fact of life in much of the world, and the need for strong defense forces continues. He stressed the importance of the discipline of military psychology, and pointed out that even with the sophistication of modern weapons, these weapons are controlled by people who still have the same limitations they have always had. Only by understanding human behavior can we hope to maximize the effectiveness of both the personnel and the equipment that comprise a strong defense force. In particular he stressed the significance of research in human factors and in training, noting that recent experiences in the Gulf war demonstrated the importance of training as a means of helping to counteract numerical superiority.

DECISION-MAKING

A new feature of IAMPS this year was the identification of a special focus topic that would provide an opportunity for a more detailed treatment of a subject of current interest in many nations. The topic selected for emphasis was *decision-making*, focusing on the development of decision support systems and training techniques for individual and group decision-making in military environments. Three invited papers on this subject were presented and are published in their entirety in this report. Two of the papers summarize recent trends in decision making research, both in Europe (the paper by Prof. Berndt Brehmer) and in the United States (by Dr. Martin Tolcott). The third paper (by Dr. Janis Cannon-Bowers et al.), describes an important ongoing program to improve tactical decision-making under stressful conditions. This is a topic of widespread interest in Europe as well as North America; the paper discusses the scientific underpinnings on which the program is based. Other presentations relevant to decision making are discussed below.

Doctor Lars Rejnus, a Senior Research Officer at FOA, discussed a decision support system called NBCBRA, which was developed for responding to nuclear/biological/chemical (NBC) attacks. This system, currently in use, serves not only as a decision support tool but also for training and for generating baseline data in various hypothetical scenarios.

Rejnus's paper focused on issues related to development of the user interface for the system. A natural language interface was developed which allows even novice users to communicate effectively with the system without extensive training. The system is queried by means of grammatically correct sentences that are constructed in part by means of sentence fragments chosen from pop-up menus. During system development substantial attention was paid to usability considerations, which were discussed. Interface features include a help system, extensive use of windows, ready availability of backup information, keystroke accelerators, legend text used in conjunction with windows, and the use of colors in geographical displays.

The next generation of this system will have a more advanced capability for generating and manipulating text. Future research will compare the current interface with other approaches now possible with graphical user interfaces (such as MS Windows).

Doctor David J. Hickson, Professor of International Management and Organization at the University of Bradford Management Center, England, presented a paper (coauthored by Sue Miller and David C. Wilson) on making top-level strategic decisions in public and private sector organizations. The paper was based in part on recently published material by the authors and others.[1,2]

Hickson's work has focused on understanding the various decision-making processes that exist in organizations; he describes three general patterns:

- *sporadic processes*, in which decision making is drawn out over a protracted period of time, with frequent delays, obstacles, and detours;
- *fluid processes*, which proceed more smoothly, quickly, and systematically; and
- *constricted processes*, which are more tightly controlled by a fewer number of participants, with less negotiating and consensus-building.

Although not done in a military context, the research can be applied to certain types of military decisions. For example, in examining organizations with similar structures to those typically found in military settings (clear-cut hierarchical control combined with decentralized tactical flexibility), Hickson notes that such organizations tend to have unduly constricted decision-making processes.

Future work will concentrate on determining what makes a decision successful — in particular, the extent to which the decision-making process itself affects the quality of the outcome, and therefore what organizations can do to improve that process. Some preliminary speculation on this issue has led to the suggestion that success depends on two kinds of factors: "launchers," which help get a decision moving in the right direction (factors such as timing, relevant experience, planning, commitment by all major participants) and "propellants," which sustain its momentum (expertise, competence, effective lines of control and responsibility).

PERSONNEL SELECTION AND CLASSIFICATION

The development and utilization of tests for determining fitness for military service and for assigning people to jobs for which they are well suited continues to be a subject of interest in many countries. This section summarizes several presentations related to this topic.

Chief Psychologist Johan Lothigius, of the Swedish National Service Administration, gave an overview of activities of the psychologists in his organization, which, as the enrollment board of the Swedish Armed Forces, is the central authority for enlistment and other personnel matters. The head office is in Karlstad, with five regional offices around the country.

Since the 13th century, Sweden has had compulsory military service for its male citizens. Currently this affects men beginning at age 18, who, after completing basic training, receive periodic refresher training and are subject to callup until age 47. The armed forces have about 40,000 permanent employees, including 16,000 career officers.

Between 50,000 and 60,000 people are enlisted annually. The psychological examinations associated with the enlistment process include the following components:

- a test of general intelligence, consisting of four subtests for logical, verbal, spatial and technical abilities;
- various biographical and attitudinal questionnaires;
- tests of mechanical abilities and telegraphy tests (not given to all conscripts); and
- an interview and assessment, which is the most important aspect of the psychological evaluation process.

The three principal factors assessed during the interview are:

- mental fitness (about 7% are rejected on this factor);
- ability to cope with the stress of war; and
- leadership ability.

The psychological tests are used, along with preference statements, by an enrollment officer who makes assignment decisions.

Post-enrollment tasks performed by Swedish military psychologists include a continuing assessment of mental fitness, and special missions in education, research and personnel selection.

Lothigius stated that there is substantial interest in the eventual use of computer-based tests, including a computerized adaptive testing system for selection and assignment. Potential tests are currently under development and evaluation.

Doctor Friedrich Steege, of the Psychology Service of the German Ministry of Defense, discussed simulation-based approaches to personnel selection and classification. A premise of these approaches is that the increased realism provided by using computers to simulate portions of the criterion task will yield improved predictive power.

Steege first discussed some constructs provided by cognitive psychology that describe underlying abilities or dimensions that many computer-based tests attempt to assess. For example, the taxonomy described in Ref. 3 lists the following constructs: spatial orientation and visualization, numerical facility, time sharing and selective attention, reaction time and choice reaction, psychomotor skills, and complex information processing. Steege noted that much work remains in developing tests and constructs that open new dimensions in both the predictor and criterion domains, and in using computers to measure abilities that cannot be assessed by paper-and-pencil tests.[4]

After mentioning several recent or ongoing programs to extend the use of computers for testing, Steege summarized recent developments in simulation-based assessment, which has been used effectively in aircraft pilot selection. He noted that the simulation approach allows the most salient features of the task to be emphasized, by eliminating other, less relevant, aspects. He then summarized the approaches to pilot selection being taken by several nations and discussed in greater detail the system being developed in the Federal Republic of Germany (FRG). This system, the "Instrument Coordination Analyzer" (ICA 90), is a computer-based battery that simulates various tasks relevant to flying. ICA 90 contains five tests that together represent a wide range of abilities considered essential to flying: psychomotor coordination, coding of information and flexibility of mental representation, anticipation of self-motion in space, capacity and strategies of information processing, and problem-solving competence. The test battery is being given to new officers having no relevant flight experience. Baseline data are being collected.

There is a great deal of current interest in Europe and North America in the use of computer-based tests to tap human abilities that cannot be adequately evaluated with traditional tests. Pilot selection is clearly one of the principal areas that is benefitting from this interest, and one in which there is substantial activity in many countries. The need to conduct comparative evaluations of various approaches being pursued is becoming increasingly clear. IAMPS is one useful forum for information exchange on this topic.

Major Maxon Mosher, of the Canadian Forces Directorate of Recruiting and Selection, discussed another approach to personnel selection — the use of interviews. Although reviews of prior research have cast doubt on the reliability and validity of employment interviews,[5,6] recent research, involving highly structured interviews, has shown considerable promise.[7] Mosher described a recent effort to examine the usefulness of structured interviews for selecting personnel for the Canadian Forces (CF). One of the factors motivating this research was recent Canadian human rights legislation, which requires all Federal employers to demonstrate the job-relevance of any information obtained during the selection process. Because job interviews are relatively straightforward to defend on this basis, they may be given greater weight in the future, especially if their predictive validity can be improved.

In general, structured interviews are based on critical incident analyses in which experts identify critical behaviors in various job-relevant situations, and interview questions are then developed that relate to those behaviors. For example, target behaviors of interest in selecting military recruits include:

- conduct (following rules, accepting criticism);
- teamwork (getting along with peers, participating in group activities); and
- coping (adapting to new or unusual situations, reacting appropriately to stress).

The research described by Mosher compared two similar interview techniques: the Situational Interview (SI) approach developed by Latham (Latham and Saari, 1984), in which applicants are asked what they would do in various job-related situations; and the Patterned Behavior Description Interview (PBDI) approach of Janz (Janz, Hellervik and Gilmore, 1986), in which the questions relate to actual past behavior. Subjects were CF recruits, interviewed prior to entering recruit training. Criterion measures were obtained from the recruit school at the end of the 10-week training program.

The results favored the PBDI, which had a significant validity coefficient ($r = .44$, $n = 65$, $p < .001$). Methodological difficulties in administering the SI may have contributed to its relatively poor showing ($r = .14$, $n = 18$), which was not statistically significant. Future training for CF Military Career Counsellors who will be assigned to recruiting duties will include an interviewing approach based on the PBDI.

Doctor Svend E. Olsen, from the Danish Defense Center for Leadership, presented results from a preliminary study to determine the characteristics of effective tank commanders with regard to "social competence," which encompasses a variety of skills and abilities related to social interaction. The factors judged especially relevant to tank commanders were: informal leadership, flexibility/adaptability, stress resistance, social situational awareness, empathy, self-confidence, and authority.

Results of this work suggested that the most effective tank commanders exhibit a "person-oriented" leadership style. Evidence also supported the hypothesis that more effective commanders possess a greater amount of self-knowledge: the self-ratings made by the more competent commanders tended to correlate higher with external ratings than the self-ratings of the less effective commanders. Self-knowledge has long been considered an important component of good leadership.[8] More generally, issues of small-group leadership, with application to both selection and training, are increasingly being seen as important areas of research in the context of a variety of military teams.

Lieutenant Colonel Sandro Tomassini, head of the Applied Psychology division for the Italian Armed Forces, presented a paper (coauthored by Lt. Col. Maurizio Laurenti) discussing the procedures for accepting applicants to the Army's Military Academy. He noted that until recently, personality and aptitude test scores were not used to reject applicants but only to group them into five general categories; their scores on other tests (general culture and mathematics) also affected their final rankings. This procedure has proven unsatisfactory because of high resignation rates and failure rates on written examinations.

A new procedure, now in effect, permits aptitude and personality scores to be used as a basis for rejection, and as a basis for irrevocable assignment to one of three career paths: various Armies, Carabinieri Army, and Logistics Corps. It is believed that this will result in a more efficient and less costly officer selection procedure.

Captain Antonio Peri, a psychiatrist in the Personnel Branch of the Italian Navy, presented findings (coauthored by Maria C. Ruffini and Andrea Citone) from a study to determine the personality factors and coping skills that are most conducive to working and living harmoniously aboard small ships for long periods of deployment. Overall, the study found morale on the ships investigated to be high, perceived discomfort level surprisingly low, and in general a quite satisfactory level of all socio-emotional factors measured.

Peri pointed out that the results may not be generalizable because of methodological problems and sample size. However, his finding of a positive correspondence between certain socio-emotional factors (morale, cohesion, interpersonal compatibility, general feelings of well-being) and performance improvement over time suggests that selection tests based on such factors may be useful in improving overall operational efficiency aboard ship.

Colonel Aurelio Pamplona and Capt. Antonio Roy Costa, psychologists in the Portuguese Army, evaluated the ability of several tests to assess an individual's capabilities for coping with stress. The purpose of the research was to improve the selection process for personnel to be assigned to Special Forces and Commando units. Three types of paper-and-pencil tests were administered: Rosenbaum's Self-Control Schedule, Levenson's Control Locus Scale, and Zung's Anxiety Self-Evaluation Scale. Results showed support for the use of these tests to predict the ability of students in commando school to deal appropriately with unexpected stressful events. In particular, measures of self-control as indicators of resourcefulness[9] were useful, suggesting that highly resourceful individuals use more effective coping methods when faced with stressful events.

Two psychologists from the Psychology Service of the Spanish Army discussed ongoing activities in their organization. Colonel Felix Utrilla Layna described the process by which approximately 200,000 inductees are tested each year, and individuals with suspected psychopathologies are identified. He noted an increasing appreciation for the important role played by military psychology in the Spanish Army. Lieutenant Colonel Pablo Lazaro Pueyo described the battery of psychological tests administered to enlistees. He emphasized those used to detect problems that would render an individual unsuitable for military service. The principal screening tool for identifying possible psychological disorders is a personality test having scales for depression, psychopathic deviation, neurosis, and extroversion. Follow-up clinical interviews are conducted for examinees scoring beyond a critical level on one or more of the scales. Disqualifying individuals for service based on this procedure has resulted in a decline in the suicide rate during the past five years.

TRAINING

Development of improved instructional methods and cost-effective simulation systems for military training is an increasingly important activity of military psychologists in many nations. Presentations on this subject are summarized in this section.

Doctor Maud Angelborg-Thanderz from the Division of Aviation Medicine at FOA reported on research aimed at determining whether, or to what extent, pilot proficiency can be maintained or rapidly reacquired when flight training is done on an intermittent basis, interspersed with relatively long periods of no training. The research was motivated by

- questions about how to train pilots who would be required to fly a variety of types of missions,

- concern about rising training costs and decreasing budgets, and
- an interest in being able to periodically use former military pilots now flying for the airline industry.

Pilots who had not flown a particular military aircraft for periods ranging from one-half year to 12 years were given an intensive simulator-based training program. This was followed by evaluations of performance and workload, both in the simulator and in actual flight. Workload measures were similar to those discussed by Svensson in his IAMPS presentation [see below] and reported elsewhere.[10] Performance in the aircraft was generally very good for these pilots. In some cases they performed better than the younger pilots with current aircraft experience, but at a cost of substantially higher workload.

The major conclusion from this work is that highly skilled pilots can recover from long periods away from the aircraft, provided they receive individualized refresher training from experienced instructors who provide consistent and frequent feedback. The work also validated the importance of flight simulators as cost-effective substitutes for actual in-flight training.

Doctor Erland Svensson of FOA's Division of Aviation Medicine presented a paper (coauthored by Maud Angelborg-Thanderz and Lennart Sjöberg) describing research at FOA to develop an index of a pilot's mental workload during a variety of missions. Such an index would be useful, among other things, as a means of evaluating decision support systems, of analyzing specific missions, and of measuring the effectiveness of training. Although substantial progress has been made in the area of workload measurement (see, for example, Refs. 11 and 12), much work remains to be done. Clearly workload is a multidimensional concept that cannot be reliably assessed by a single measure.[13] The three general approaches to workload assessment — subjective ratings, objective performance measures, and physiological measures — each have their own strengths and weaknesses. The work described by Svensson attempts to improve the reliability and validity of subjective ratings by combining them with physiological measures into a single index.

Based on results from flight tests, a model was developed that incorporates a set of psychological variables (subjective ratings of difficulty, risk, effort, mood, etc.) and psychophysiological variables (adrenaline and noradrenaline excretion levels). These variables were combined into a Workload Index (WI), which characterizes high workload as consisting of increased tension or mental stress, increased effort (psychological and physiological), and increased energy mobilization followed by fatigue. The WI was validated in simulated and actual flight tests. Results showed it to be a sensitive indicator of experience level, with workload dropping substantially as training progressed.

Svensson reported that work is now underway to compare the WI with other indices, such as the NASA Task Load Index and other indices.

Lieutenant Colonel Jean-Pierre Pauchard, a psychiatrist in the Swiss Army, presented the results of preliminary research into the use of a novel method for leadership training that focuses on the causes of poor interpersonal relationships. This technique, called the Balint group method (named after a 20th-century Hungarian psychoanalyst) was first tested in a military setting by the Swiss in 1985. It is now being used for officer training in the Military Leadership Training School in Zurich. The goal of the training is to teach leaders how to detect, understand, and repair disturbed relationships (with subordinates, superiors, or colleagues) that can undermine the leader's effectiveness regardless of whether they involve him directly or only indirectly.

The essence of the Balint method is the presentation of a case history in a group setting. As adapted for military leadership training, the group consists of a professionally-trained leader and co-leader along with 10-12 trainees who listen to one of their group discuss a problem relationship with which he or she is personally familiar. Members of the group ask questions and then engage in a brainstorming session in which observations and personal reactions to the story are discussed; explanations or advice are prohibited. This is followed by a period of general discussion. The aim of the session is not to try to solve the problem, but rather to examine the conditions that led to it.

Pauchard reported generally good success with this technique as a way to teach military leaders how to diagnose the cause of difficulties with relationships. One of its principal benefits is that it helps officers, who may be ignorant of (or antagonistic toward) psychology, to realize that they can use their innate knowledge of people and their observational skills in ways that will help them understand and eliminate many interpersonal conflicts. He cautioned that the method should be used only by instructors with a psychoanalytic background as well as experience in group dynamics.

First Lieutenant Jürg Stadelmann, a psychologist in the Swiss Army, described a recent project aimed at counteracting a commonly observed phenomenon in many countries, namely a decline in motivational level of conscripts during basic training. He suggested that this problem has increased in the last few decades as the result of a growing disparity between life in the army and the conscripts' civilian family life (which is freer and less hierarchical than in former generations). The approach taken by this research to enhance motivation was to improve the leadership skills of the non-commissioned officers (corporals) who interact most directly with the conscripts.

Leadership training consisted primarily of efforts to improve communication skills to enable the corporals to cope with difficult situations more effectively without resorting to formal authority. Two psychologists supervised the corporals in one company during an eight-week period of basic training, providing suggestions and guidance in daily meetings. The most important training tool was the use of videotapes which enabled them to provide direct and concrete feedback. Results showed that conscripts in the experimental group developed a more positive attitude about their training and about the army in general.

OTHER TOPICS

This section contains brief descriptions of presentations given on additional subjects, including stress, suicide, and factors affecting retention.

Doctor Herbert Aschenbrenner, a psychologist in Germany's Armed Forces Personnel Office, reported on the planned activities of a newly formed NATO Research Study Group (RSG-22, "Psychological Support for Military Personnel"), which he chairs. This study group evolved from a related group whose charter was to develop measurement methods and selection criteria for stress-resistance.[14] RSG-22 will develop a multinational collaborative research program related to the management of combat stress in the military. In general, this will involve identifying the kinds of situations warranting psychological support, and determining the nature of that support. Topics to be studied include training in stress coping techniques and first aid for combat stress reactions (including self-help, buddy aid, and interventions by the military leader). Work is still in the preliminary stage; it is anticipated that reports of progress will be given at future IAMPS meetings.

Lieutenant Juan José Cerezo Ureta, a psychologist in the Spanish Ministry of Defense (Air Force), presented a general tutorial on the measurement of arousal (physiological, biochemical and psychometric methods), on the relationships between arousal and performance in the context of military operations, and on techniques for regulation of arousal levels (relaxation and activation strategies). He noted that work is in progress to train pilots in the reduction of arousal (both somatic and cognitive relaxation techniques), and that preliminary results are encouraging.

Lieutenant Colonel Colman Goggin presented the results of an in-depth analysis of suicides in the Irish Permanent Defense Force (PDF) during the period 1974-1990. The average yearly suicide rate in the PDF was almost identical to the estimated national rate (approximately 13 per 100,000). Alcohol abuse, depression, and the copy-cat phenomenon were identified as major precipitating factors. The principal recommendations emerging from the study were the establishment of a comprehensive and multidisciplinary support network (involving medical officers, chaplains, and personnel support services), and the development of alcohol abuse, depression, and stress intervention programs.

Doctor Sarah Smith, a psychologist in the U.K.'s Army Personnel Research Establishment, reported on the results from a study of officers who chose to leave the Army prematurely, i.e., before the normal retirement date. Findings from an exit questionnaire that has been administered routinely since 1986 indicated no single major factor accounting for the departures; rather, a combination of reasons was most often cited. Insufficient job satisfaction, inadequate career development programs, and various lifestyle and domestic considerations were the most commonly cited concerns. Overall pay level was not identified as a major determining factor.

Doctor Milton Katz presented an overview of the mission and functions of the U.S. Army Research Institute (ARI) for the Behavioral and Social Sciences, with emphasis on ARI's European Science Coordination Office (ARIESCO), which he heads. In addition to performing a liaison function between ARI and military researchers outside the U.S., ARIESCO funds basic research that complements ongoing U.S. programs. Emphasis is placed on efforts: that reflect different scientific approaches to the study of certain topics, and that would be difficult or impossible to conduct in the U.S. for various reasons. Examples of topics currently supported include performance under stress, terrorism, unit cohesion, and courageous behavior.

CLOSING REMARKS

As can be seen from this summary of the symposium, IAMPS presentations typically cover a wide range of subjects. These include controlled laboratory experiments, field studies, preliminary inquiries, analyses, and tutorials. This variety reflects the diversity of the activities across nations and the amount of resources available for research and development. The value of this symposium derives not just from the exchange of scientific results but also from a sharing of experiences and lessons learned from diverse societies and cultures, most of which are facing similar challenges in an era of rapidly-changing geopolitics.

I believe that military psychology is destined to play an increasingly important role, especially in those countries facing a period of force downsizing and defense budget reductions. These factors will result in an increased emphasis on the efficient use of manpower resources, on the cost-effectiveness of training, and on the operability and maintainability of weapons systems. I am confident that IAMPS will continue to be a useful and informative forum in the coming years of turbulence and change in the armed forces of many nations.

With regard to future IAMPS agendas, the idea of identifying a special focus topic each year was well-received by the symposium participants and will be continued.

REFERENCES

1. D.J. Hickson and S. Miller, "Concepts of Decision: Making and Implementing Strategic Decisions in Organizations," in *Leadership and Decision Making* (Cambridge University Press, Cambridge U.K., 1991), F. Heller (ed).

2. D.J. Hickson, R.J. Butler, D. Cray, G.R. Mallory, and D.C. Wilson, "Decision and Organization" *Public Administration 67*, 373-390 (1989).

3. C.J. Martin, *A Review of the Reliability and Validty of Computerized Tests Among RSG 15 NATO Countries.* (Draft paper) Department of the Navy, Washington, DC.

4. F.W. Steege and W. Fritscher, "Psychological Assessment and Military Personnel Management," in *Handbook of Military Psychology.* Wiley Chichester, (1991) R. Gal and A.D. Mangelsdorff, eds.

5. E.C. Webster, *Decision Making in the Employment Interview* (Eagle Publishing Co., Montreal, 1964).

6. E.C. Webster, *The Employment Interview: A Social Judgment Process* (SIP Publications, Schomberg, Ontario, 1982).

7. M.M. Harris, "Reconsidering the Employment Interview: A Review of Recent Literature and Suggestions for Future Research" *Personnel Psychology 42*, 691-726 (1989)

8. G.C. Homans, *The Human Group.* (Harcourt, Brace & Co., New York, 1950).

9. M. Rosenbaum, "A Schedule for Assessing Self-control Behaviors. Preliminary Findings" *Behavior Therapy, 11,* 109-121 (1980).

10. E. Svensson, M. Angelborg-Thanderz, L. Sjöberg, and M. Gillberg, "Military Flight Experience and Sympatho-adrenal Activity," *Aviation, Space and Environmental Medicine, 59,* 411-416 (1988).

11. N. Moray, N. (ed.) *Mental workload: Its Theory and Measurement.* (Plenum, New York, 1979).

12. P.A. Hancock and N. Meshkati, (Eds.) *Human Mental Workload*, (Elsevier Science Publishers, Amsterdam, 1988).

13. D. Gopher and E. Donchin, "Workload - An Examination of the Concept," in *Handbook of Perception and Human Performance, Volume II*, (John Wiley & Sons, Inc. New York,1986). K.R. Boff, L. Kaufman and J.P. Thomas, (eds.).

14. H. Aschenbrenner and D. Wiegand, *Proceedings of the Workshop on Psychological Fitness* (Lahr, Germany, 28-29 Sept. 1988). *NATO Document AC243/8-DS/A/DR(89), p. 107.*

ADDITIONAL REFERENCES

K.A. Holroyd and R.S. Lazarus, 1983. "Stress, Coping and Somatic Adaptation," in *Handbook of Stress* (New York: Macmillan, 1983), Goldberger and Brenitz (Eds.).

T. Janz, L. Hellervik, and D.C. Gilmore, *Behavior Description Interviewing* Allyn and Bacon, Inc., (M.A. Newton; 1986).

G.P. Latham and L.M. Saari, "Do People Do What They Say? Further Studies on the Situational Interview," *Journal of Applied Psychology*, 69, 569-573 (1984).

European Approaches to the Study of Decision Making: Focus on Process

by Berndt Brehmer, Professor of Psychology, Uppsala University, Uppsala, Sweden.

ABSTRACT

European research on decision making is distinguished by a focus on process. This is manifested both in research on the process of decision making, i.e., on what goes on between the moment a decision problem occurs and that when the decision maker arrives at a decision, and in research on how people control processes. The former line of research tends to view decision making as a problem solving process and analyzes it into a series of stages, and it relies on "think aloud" protocols. Two examples are discussed: Huber's operator model which analyzes decision making in terms of a series of short-term memory contents transformed by operators, and Montgomery's dominance structuring model which discusses decision from the point of view of its function: to provide stable intentions for action by searching for (and sometimes creating) a dominance structure. Examples of the latter approach are the German tradition of research on complex problem solving and research on dynamic decision making. This research relies on computer simulated microworlds which are used to study how people achieve control over complex systems. The research is discussed in terms of two different approaches: the individual differences approach and the experimental approach, and results from these two approaches are briefly reviewed. The results from the individual differences approach show that there are considerable inter-individual differences in decision making performance, that it has been difficult to find correlations between performance and standard psychological tests, and that there are characteristic errors, the "pathologies of decision making", that unsuccessful decision makers exhibit. Results from the experimental approach show that certain system characteristics, such as feedback delays and side effects of decisions, have strong negative effects of decision performance.

INTRODUCTION

Decision research is a very lively area in European psychology, and it covers, in one form or other, most aspects of the topic. The conferences on Subjective Probability, Utility and Decision Making (SPUDM) have been held biannually in Europe since the early sixties (the 13th of these conference was held in Fribourg this year.) The proceedings from these conferences (see Brehmer, Jungermann, Lourens & Sevon, 1986, Rohrmann, Beach, Vlek & Watson, 1989, and Borcherding, Larichev & Messick, 1990, for the proceedings of the 10th, 11th and 12th SPUDM conferences, respectively) give a good overview of European decision research (and some American research, since the SPUDM conferences are international conferences).

However, even though European research covers many topics that are common to those of interest to American researchers, it also has some distinctive features. Thus, it differs from American research in being relatively less interested in normative issues, and in being more interested in process. The latter interest is manifested in two ways.

First there is a focus on the *process of arriving at a decision.* This approach involves using "think aloud" methods to elucidate what goes on between the moment a decision problem is presented and that when a decision is reached. The general focus here may be said to be the view of *decision making as a problem solving process.* A special group of European researchers, united in a group called the European Group for Process Tracing Studies of Decision Making, have met regularly since 1982 to develop this approach. A recent book edited by Montgomery and Svenson (1989) gives a good overview of the results obtained.

Second, interest in process is manifested in the research on *the role of decision making in the control of processes*, exemplified both in the German tradition of research on *"komplexes Problemlösen"* and in our own work on *dynamic decision making.* This research is characterized by the use of a new form of research methodology, that of computer simulated microworlds, i.e., complex computer simulations that are used to study how people cope with processes.

It is, of course, possible to find examples of both of these approaches in the United States also, but they are hardly as prominent features of American decision research as of the European research.

This paper will give examples of both of these approaches, starting with research on decision making as a process.

DECISION MAKING AS A PROCESS

Research on decision making has always had a strong interest in normative issues. There are at least two reasons for this. The first is that decision researchers have had a strong commitment to applications, and to improve decision making. This requires a model of what a good decision is, so that it is possible to evaluate the decisions made both when the decision maker is unaided and when he has access to some decision aid. Without such a normative model it would obviously be impossible to determine if the decision maker actually needed decision aids.

The second reason is theoretical. Here, research on decision making has been seen as an opportunity for investigating a problem of eternal interest: viz., the extent to which man is *rational.*

Statistical decision theory using some variety of the expected value criterion has served to define a good decision for both these lines of research. That is, a good decision has been defined as the choice of that alternative which maximizes expected value. This criterion has been questioned for some time (see, for example, Einhorn & Hogarth, 1981), and there seems to be less interest in normative issues in decision making recently.

For the purpose of ascertaining the rationality of decisions, it is sufficient to study input-output relations, and to compare the decisions made by a group of subjects with some normatively prescribed ideal decision. However, interesting as these normative analyses may be, they nevertheless fail to elucidate the psychological problem of understanding how decisions come about. For this, we also want to know what goes on in the decision maker's mind from the presentation of a decision problem to the decision. This has been the focus of European research of the process of decision making.

One starting point for a process analysis is to identify a series of *stages* that a decision process may go through. Many researchers have tried to identify a number of such stages. One stage might be identification of a problem, a second the generation of options, a third the evaluation of options, and a fourth choosing an option. Such stages serve to identify the components of the decision problem and to define such a problem, but they do not necessarily identify the stages that a decision maker might actually go through. However, such a stage analysis brings out the similarities between decision making and problem solving. It is therefore not surprising that these analyses of decision making have resulted in attempts to incorporate approaches that have proved their worth in the study of problem solving also into the field of decision making. This has involved taking over both methods and theory from problem solving research. The most important method is undoubtedly analysis of verbal protocols obtained by "think aloud" methods. In experiments using this method, the subjects are encouraged to report everything that comes into their minds, and the experimenters then use these utterances to reconstruct the subjects' cognitive processes. This can be done at various levels of detail. Thus, one approach involves constructing a protocol which shows the subject's path from the initial state to the decision in terms of a set of short-term memory contents. The researcher then tries to find a minimum set of rules that will generate the content of the protocol. Alternatively, the researcher may use the protocol as a whole to obtain more general information about the decision process.

The use of such verbal protocols is, of course, fraught with difficulties and controversy. A good discussion of the problems and various ways in which they might be solved is given by Svenson (1989). A more general discussion of protocol analysis is given in Ericsson and Simon (1984).

Protocol analysis is, of course, not the only method used for process analysis. There is also a variety of more behaviorally oriented methods, e.g., information boards, which make it possible to make inferences about cognitive processes involved in decision making from information about how and in what order subjects select information (see Payne, 1976).

Protocol analysis is not limited to experimental data. It is also possible to use various forms of documents as a basis for analysis of decision processes. For example, Biel and Montgomery (1989) applied protocol analysis to transcripts of speeches of the Swedish minister of energy to demonstrate the generality of Montgomery's dominance structuring model, of which more will be said later. Gallhofer and Sarris (e.g., 1989) have analyzed Dutch foreign policy decisions in a series of papers, and demonstrated that many of the phenomena uncovered in experimental research on decision making can be found in these documents, thus supporting the generality of the experimental findings. Finally, Crozier (1989) has analyzed economic and political decisions, e.g., the De Lorean affair, i.e., the decision to fund the De Lorean sports cars factory in Belfast. I am convinced that we will see more of these kinds of applications in the future.

However, decision researchers have not only taken over methods from problem solving research. They have also tried to apply some of the theoretical insights from research in that area. One such attempt is Huber's operator approach to decision making.

Huber's Operator Approach

In his book *Entscheiden als Problemlösen (Decision making as problem solving)* (1982) and in many subsequent publications, Oswald Huber, now at the University of Fribourg, has reported on the progress towards developing a model of decision making based on the Newell and Simon (1972) approach to problem solving.

As indicated by the title of his book, Huber sees decision making as a form of problem solving. He points out the similarities between the two:

> "*The decision process starts with an initial situation (initial state) in which, for example, the decision maker is confronted with a set of alternatives to choose from. There is a desired goal state situation (goal state), in which exactly one alternative has been chosen (in a competent manner). The problem solver does not know in advance how to transform the initial state into the goal state (provided the specific decision is not a routine task).*" (Huber, 1989, p. 4, italics in the original)

In short, the decision maker's problem is to find a way to transform the initial state into the goal state. This formulation agrees closely with the view of decision making in Newell and Simon (1972).

Huber calls his model the *operator model*, and it describes decision making as a series of short-term memory contents, changed by operators. These operators transform some initial state into a goal state. Specifically, the initial state together with the set of operators make it possible to construct a *decision space*, similar to the problem space in the Newell and Simon approach. Finding a decision then becomes a matter of searching through this decision space until an acceptable alternative has been found.

Huber then goes on to search for the operators that are necessary to effect the kinds of transformations that are observed in think aloud protocols from decision making experiments, which are often rather traditional forms of experiments involving lotteries. He finds such operators at two levels: *elementary operators* which change one state into another, and *complex operators* which comprise a series of elementary operators.

The *elementary operators* in Huber's model are the following:

- *Evaluation operators* (EVAL + and EVAL -) which perform simple evaluations of objects.
- *Selection operators* which select objects with the highest value (MAX), the lowest value (MIN) or equal value (EQUAL).
- *Concatenation operators*, that combine elementary contents into some more complex wholes, e.g. summation.
- *Difference operators*, i.e., operators which produce a difference as the result of a comparison.
- *Criterion operators*, which define the level of acceptance.
- *Disarmable operators*, which test if an aspect of an object can be transformed into a more fitting one.

The complex operators are sequences of elementary operators. The complex heuristics in Huber's model are basically different forms of heuristics. Thus, there are:

- *Subheuristics*, consisting of decision rules, such as dominance or lexicographic ordering,
- *Decision heuristics* for evaluating information, and
- *Structuring plans* which transform a decision problem into some more useful form for further processing.

Thus, Huber's model, like the original Newell and Simon model for problem solving, is built up by a small set of simple, but powerful operators. The undeniable complexity of decision behavior comes about from the repeated application of these simple operators to complex short-term memory contents.

Note that the elementary operators in Huber's model are quite general and in no way specific to decision making. What is specific to decision making is found in the complex operators. No special cognitive mechanisms are needed and decision making is thus integrated into cognitive psychology.

Space limitations prevent us from going into more detail, and our account here can, of course, only give the general flavor of Huber's model. The actual application necessarily leads to very detailed analyses, and it is best understood by going through actual protocols of decision behavior.

It is still too early to decide how successful this approach will be, but the results presented in some of Huber's later papers appear promising both for the more traditional lottery type decision tasks and for sequential decision tasks (see, e.g., Huber, 1989).

Montgomery's Dominance Structuring Model

Huber's work exemplifies the use of protocol on a very detailed level of analysis and a method closely allied to that of Newell and Simon (1972). Montgomery's work (Montgomery, 1983, 1989), on the other hand, is an example of a more molar form of theorizing which stops at a level corresponding to that of complex operators in Huber's model although Montgomery's concepts are less formally defined than Huber's, so it is not possible to make any detailed comparisons between the two models. Moreover, Montgomery uses protocols in a freer way, obviously seeing them as a source of data to be mined for interesting nuggets of information, rather than as the starting point for a detailed sequence of short-term memory states and operators.

Montgomery's model is called the *dominance structuring model.* Its basic idea is that we decide on those alternatives for which we can find good arguments, i.e., arguments that will hold even if we get new information, and even if we have to defend our decisions against an opponent with a different view. This is a good principle because it makes it possible for us to maintain our intention to act, and it thus makes it more likely that we will eventually do something and not be lost in thought. The function of decisions, then, is to prepare for action, and to make sure that actions are indeed carried out. Referring to Ajzen and Fishbein (1980) he sees intentions as the guiding forces behind actions, so it becomes very important to be able to maintain one's intentions. The problem of actually implementing our decisions is a problem concerning many European researchers, see Kuhl (1986) and Sjöberg (1980). Montgomery's work suggests that the implementation and action problems may shape the cognitive processes involved in decision making.

The best way of insuring a viable decision meeting the above criteria is to find a *dominant* alternative, i.e., one that is better than all other alternatives on all dimensions. Such alternatives are likely to survive the criticism from others and self better than other alternatives. Consequently, a dominant alternative is likely to be a better guide for action than an otherwise attractive, but nondominant, alternative. However, the world does not always provide such dominant alternatives. Hence, the decision maker must structure the information available to him in such a way that a dominant alternative emerges. Montgomery's model describes how this may occur. Specifically, his model assumes a series of four steps or stages.

The first stage is called the *pre-editing stage*. Here, the decision maker selects the alternatives and attributes to be included in the representation of the problem. In this stage, attributes that are seen as unimportant are discarded, as are alternatives that are considered very unlikely to become dominant. In particular, alternatives that are seen as very unattractive are eliminated.

The second stage is concerned with *finding a promising alternative*, i.e., in this stage the decision maker searches for an alternative that is likely to become dominant. For example, an alternative that is more attractive than other alternatives on an important attribute may be chosen as a promising candidate.

The third stage is called the *dominance testing stage*. Here, the decision maker tests whether a promising alternative is indeed dominant over the others. Most important, he looks for possible disadvantages that the promising alterative may have compared to other alternatives, or compared to some absolute criterion values (e.g., cost). These tests could be more or less systematic. If the promising alternative is found to be dominant, the process stops here, and a decision is made.

If the decision maker finds that the promising alternative violates a dominance structure, he may enter the *dominance structuring stage*. The goal of this stage is to restructure the information in such a way that a dominance structure is achieved. To achieve this, the decision maker attempts to neutralize or counterbalance the disadvantage(s) found for the promising alternative. There is a variety of ways of doing this.

1. *De-emphasizing the disadvantages:* e.g., arguing that there is only a very low probability that a given disadvantage will materialize, or a high probability that it can be controlled or avoided in some way, or other. Alternatively, the decision maker could argue that the disadvantage is not very great compared to those of other alternatives, or compared to some criterion value.
2. *Bolstering the advantages of the promising alternative.* This means that the decision maker enhances the importance of some attribute, e.g., by creating very vivid images of what this aspect stands for. As a consequence, the less favorable dimensions may be seen as less important.
3. *Cancellation,* i.e, the decision maker makes a trade-off with some other aspect or attribute of the problem.
4. *Collapsing,* i.e., two or more attributes are combined into a more comprehensive attribute.

If the decision maker fails to find a dominant alternative using one or more of these four strategies, he may go back to some earlier stage, e.g., selecting a new promising alternative for dominance structuring, or he may postpone the decision.

This may be seen as a recipe for irrational decision making. There are limits to the extent to which the process will go astray, however. These are set by the decision maker's *knowledge* about the decision problem. Thus, the best way of aiding decision making would be to provide the knowledge that the decision maker needs to select a useful course of action, rather than to provide procedures in the way many classical decision aids do.

Montgomery's model is one of the few models of decision making that view decisions in terms of their function for behavior. It allows us to understand some of the peculiarities of decision making from the perspective of its function: that of preparing for action by creating a stable intention. As such, it provides a more psychologically credible account of the decision process than the models which consider decision making in isolation.

We will meet the focus on action also in our discussion of the next European approach, that of dynamic decision making.

DYNAMIC DECISION MAKING

It is easiest to introduce dynamic decision making by means of an example. Consider the task facing a fire chief charged with the problem of extinguishing forest fires. He receives information about fires and their location from a spotter plane. On the basis of these reports, he issues orders to his fire fighting units, sending them to the location of the fire. The units then report back to him about their location and activities, and on the basis of this information, the fire chief sends out new orders to his units until the fire has been extinguished.

The fire chief's task is an example of a dynamic decision task. It has the following four characteristics:

1. *A series of decisions is needed.* That is, the decision maker cannot reach his or her goal by means of a single decision.
2. *The decisions are not independent,* i.e., current decisions are constrained by earlier decisions and do in turn constrain later ones. Therefore, it is not sufficient to consider only how the current decision may solve the problem at hand, but also how it will affect one's ability to cope with later decision problems.
3. *The state of the world changes during the decision process,* both autonomously, and as a consequence of the decision maker's actions. This means that the decision maker faces a difficult problem: to decide how much of an effect is due to his own actions and how much is due to the inherent processes in the system with which the decision maker is concerned. Anyone wanting to have practical examples of these difficulties should study the praise and blame that the government and the opposition respectively will apportion to the government for good versus bad outcomes of governmental decisions.
4. *The decisions are made in real time.* In these kinds of tasks, then, it is not sufficient to make the correct decisions and to make them in the correct order, they also have to be made at the correct moment in time. However, the real time demands also introduce some other important features of dynamic tasks.

First, it means that the decision maker cannot always control when, or how often, he has to make decisions. This makes decision making in dynamic tasks inherently stressful. Consequently, the decision maker must try to regulate the demands that the decision problem makes upon him. For example, a process operator may choose to run the process at a level that is below the optimal one, for the simple reason that running the process at a more optimal level leads to too unacceptable demands, requiring too many interventions on his part. Moreover, the operator may feel uncertain of whether he or she will actually be able to cope with the problems that may occur when the process is run at this level. To choose an appropriate level of control is an important aspect of skill in dynamic decision making as well as a powerful motive for actually learning to control the process.

The second consequence is that the problem in dynamic decision making must be seen as the problem of finding a way to use one process to control another process. Thus, in these kinds of tasks, not only the system to be controlled but also the means for control must be seen as processes. In the fire fighting example above, both the fire, and the fire fighting activity, must be seen as processes. The problem for the fire chief is to find a way of using the latter process to control the former. The nature of this problem is illustrated in Figure 1 for the simplest case, that of a uniform forest. This figure shows that the burning area increases linearly over time. The slope shows the speed with which the fire spreads. This parameter is obviously dependent upon the strength of the prevailing wind. The fire fighting process is also a linear process, where the slope is dependent upon the efficiency of the fire fighting units and the intercept upon the number of fire fighting units. When the slope of the fire fighting function is less than that for the fire function, it is only possible to extinguish the fire if the fire fighting units reach the fire before the fire function intersects the fire fighting function. This means that in order to extinguish the fire, the fire chief must respond quickly and massively to a fire, sending out at least as many fire fighting units as will be needed to cover the burning area when the fire fighting units are in position, which is, of course, bigger than it was at the moment of decision. For this, the fire chief needs some elementary model of fires and fire fighting, similar to that illustrated in Fig. 1.

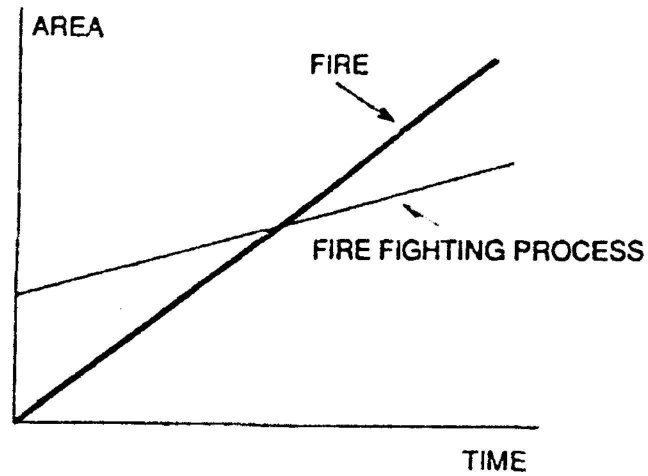

Fig. 1. The relation between the process to be controlled (the fire) and the process used for control (the fire fighting process) for the case of uniform forest.

A third consequence is that it is often necessary to take into account the various *time scales* in the task. Thus, in the fire fighting task, there are at least two time scales: that for the individual fire fighting unit which has to do with the time required for the unit to take action in the face of fire, and that for the fire chief, which has to do with the coordination of resources. Obviously, the fire chief operates in a longer time scale because it will take time for the units to move to whatever positions he sends them. Consequently, he needs to have a longer time horizon than the unit commander who only acts upon the fire that he sees. Such a set of time scales is a characteristic feature of all dynamic tasks. For example, A. Brehmer (1989) has analyzed intensive care in terms of three time scales. Here, the doctor operates in the longest time scale, setting up a treatment plan for the patient and monitoring him or her every 24 hr. The nurse carries out the plan, and sees the patient every 6 hr (when new test results arrive), deciding on possible adjustments in the implementation of the plan. Finally, the nurse's aid operates in the shortest time scale, being in charge of the minute-to-minute supervision of the state of the patient. Brehmer (1988) has analyzed the relation between different levels of command in military contexts in the same terms. Obviously, a dynamic task can be controlled only if all the important time scales in the task are supervised adequately. If only the shorter time scale is considered, the decision maker will be subject to surprises from events taking place in the longer time scale, and if the shorter time scales are neglected, nothing will be done.

It is easy to find examples of dynamic tasks of the kind discussed here. They are found in process control in industry, in patient management in hospitals, in running a company, or when fighting a battle. Despite this, dynamic tasks have not received very much study from psychologists. There are at least two reasons for this. The first is that the standard normative theory of decision making cannot be applied to these tasks. Indeed, it is often impossible to calculate the optimal decisions for these tasks from any theory (Rapoport, 1975). Consequently, they are not well suited to the research strategy of comparing actual decision behavior to optimal decision behavior which has guided so much of psychological research on decision making. Second, dynamic decision making cannot be studied with the simple paper and pencil methods characteristic of laboratory studies of decision making. Such methods do not make it possible to simulate the *interactive* nature of the decision process, i.e., the fact that the state of the task depends both on the subject's actions and on the inherent processes in the system to be controlled. Consider, for example, the practical problems involved in simulating the fire chief's decision problems using paper and pencil methods. It became feasible to study dynamic decision making only after computers had become a standard tool in the psychological laboratory, making it possible to design computer simulated microworlds with which the subject could interact in real time. We now turn to this technique.

Computer-Simulated Microworlds

A microworld is a simulation of some real system, such as a small town. In experiments with such microworlds, the focus is upon how subjects learn to control the microworld, i.e., whether he manages, either to bring it from one state to another state, or to keep it in some designated state. Such microworlds are not designed to be perfect replicas of real systems. Instead, they simulate only their most important and characteristic features, in about the same way that a portrait made from a wood cut simulates the important features of a person but leaves out the detail (Dörner, 1991).

Specifically, microworlds incorporate three important features of real world systems: complexity, dynamics and opaqueness. Thus, they are *complex* in that they require the subject to consider may different, and perhaps conflicting, goals, and different ways in which these goals can be achieved, as well as the possible side effects that stem from various courses of action. They are *dynamic* in the sense that they change both as a function of the inherent processes in the system and of the actions taken by the decision maker. Finally, they are *opaque* in the sense that they do reveal all of their features to the subject directly. Thus, the subject has to search for information, and to form and test hypotheses about the state of the microworld and its relations.

Different microworlds differ with respect to the extent to which they emphasize one or the other of these characteristics. For example, Lohhausen and Moro (see below) emphasize complexity and opaqueness, while DESSY and D^3 Fire, which were designed in Uppsala, emphasize dynamics.

Lohhausen is a microworld designed by Dörner and his associates (Dörner, Kreuzig, Reither & Stäudel, 1983). It simulates the conditions in a small German town. In experiments with this system, the subject is asked to assume the position of mayor and to run the town with dictatorial powers for a ten year period with the general task of caring for the welfare of the people of Lohhausen. The definition of welfare is left to the subject, as is the choice of courses of action. The subject has many action possibilities to choose from; there are more than 100,000 possible courses of action in Lohhausen.

Moro is another microworld designed by Dörner and coworkers (Dörner, Stäudel & Strohschneider, 1986). Moro simulates the ecological, economic and cultural conditions of the moros, a tribe living in South Sahara. The subject is asked to serve as a development worker, again with dictatorial powers, for a 20 - 30 year period, with the general goal of caring for the welfare of the moros. *Moro* makes very much the same demands as *Lohhausen* but is somewhat simpler; there are only about 1/10th of the action possibilities in Moro compared to Lohhausen.

Both Lohhausen and Moro are opaque in the sense that the subject is not informed about all aspects of the current state of the system. To find out what they need to know, the subjects must ask for information that they think they will need. For example, in an experiment with Moro, the subject may be told at the beginning of a trial that the moros complain that there is too little water for their cattle. The subject may decide to do something about this, e.g., by increasing the number of wells, but before doing so, he or she should find out what the current number of wells is, the ground water level, the number of cattle, the area of land that is currently irrigated, and so on. The experimenter who manages the computer simulation will inform the subject of this, but he will only give information about the aspects that the subject asks about.

Figure 2 shows DESSY, a microworld designed by Brehmer and Allard 1990). It requires the subject to function as a fire chief, charged with the task of extinguishing forest fires following the general scenario described in the example used to introduce dynamic decision making.

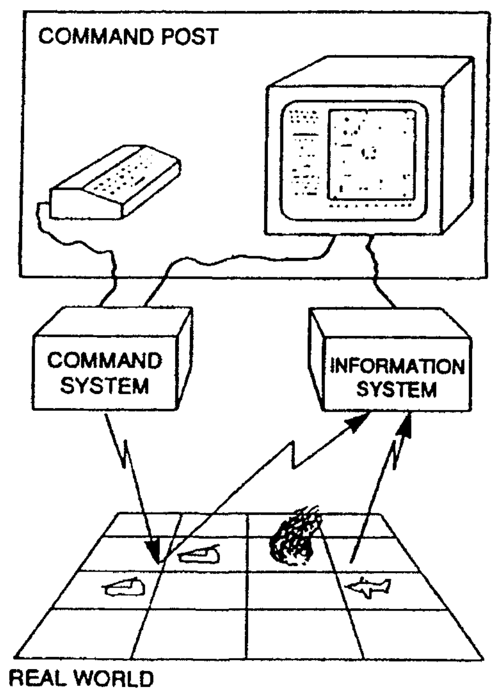

Fig. 2. The structure of the fire fighting version of DESSY.

Figure 3 finally shows D³ FIRE (Svenmarck & Brehmer, 1991) which is a version of DESSY used to study distributed decision making in dynamic contexts. It differs from DESSY in that the task is represented at the level of the individual fire fighting unit. Each subject has a limited "window" on the process, and the task for the four subjects is to coordinate their efforts so that they extinguish the fire.

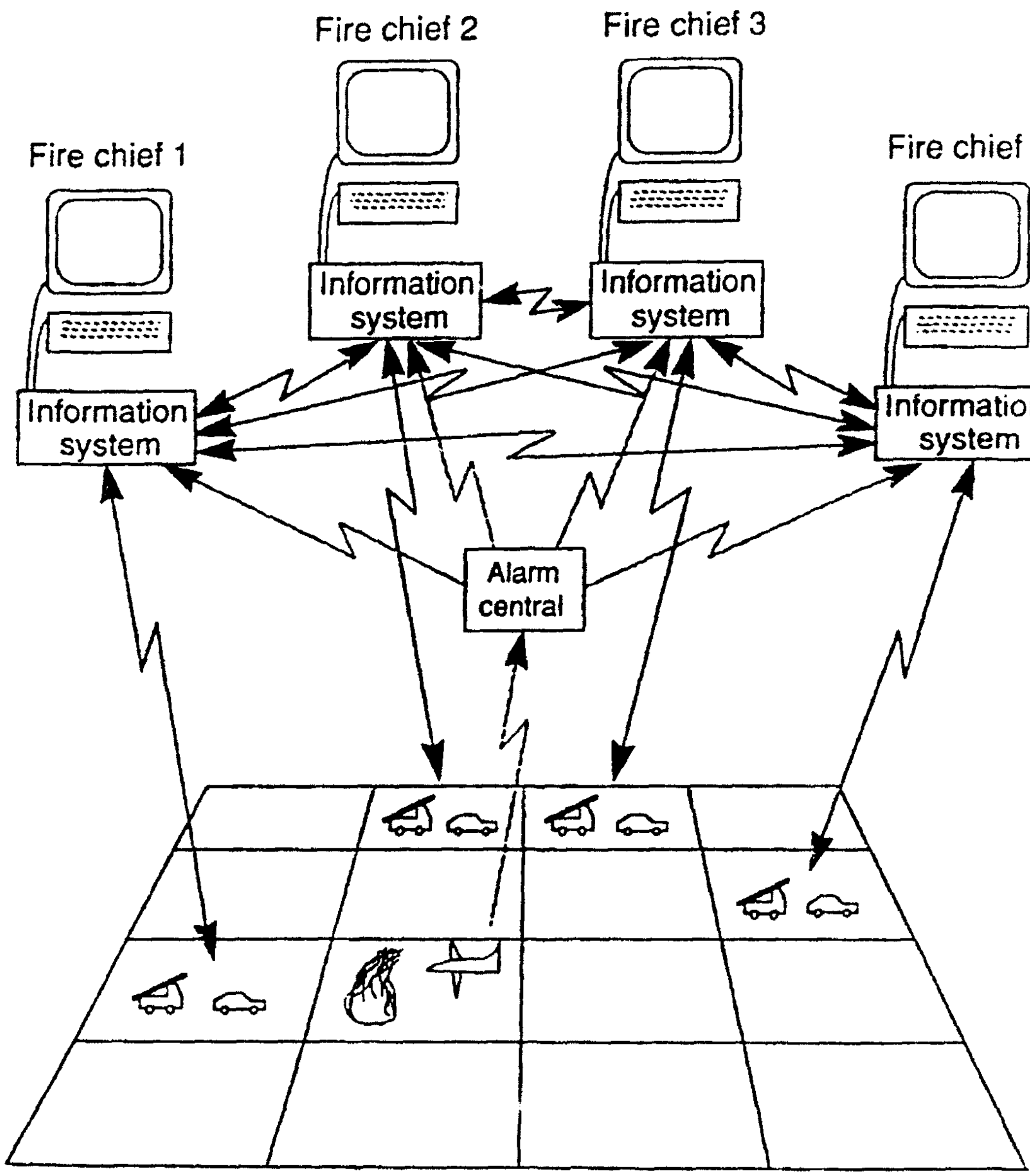

Fig. 3. The structure of D³ FIRE (Svenmarck & Brehmer, 1991).

With this brief description of the nature of some representative microworlds, we now turn to an overview of results obtained with them.

OVERVIEW OF RESULTS

Experimentation with microworlds differs in important respect from traditional psychological experiments (Brehmer & Dörner, in press). First, in experiments with microworlds, the actual problems encountered by the subject cannot be determined by the experimenter; because of the dynamic nature of these tasks, they will be a joint function of the characteristics of the system and the behavior of the subject.

Second, the independent variables cannot be defined in terms of discrete stimuli or events; they must be defined in terms of system characteristics, such as complexity, feedback delays, and the like.

Third, the dependent variables studied in microworld experiments cannot be defined in terms of discrete responses, because there it is not possible to find regularities at this level; when the sequence of problems is determined by the subject's behavior, the actual responses will, of course, differ from subject to subject. Instead, the experimenter will have to look for regularities on higher levels such as the level of strategies and tactics of the subjects. Finding these strategies and tactics is no easy task, and requires very careful data reduction. Indeed, it may be impossible to find the strategies except by simulation, i.e., a number of "ideal strategies" are defined, and the subject's behavior is then compared to these strategies by determining which of these ideal strategies account best for the subject's actual behavior (see Reichert & Dörner, 1987, for an example).

Research with microworlds has followed two different approaches: the individual differences approach and the experimental approach. We discuss each in turn.

THE INDIVIDUAL DIFFERENCES APPROACH

The individual differences approach is typical of the German work in the tradition of research on "komplexes Problemlösen" of Dörner and his associates. It involves running a group of subjects through the same microworld, and then dividing them up into extreme groups according to performance. These extreme groups are then compared with respect to their behavior in the simulation or with respect to their scores on various psychological tests.

Research following this approach has been directed at two problems: that of predicting performance in dynamic decision tasks and that of analyzing the demands made by such tasks.

Prediction of Performance

The results with respect to this goal have been disappointing; few stable correlations have been found between microworld performance and psychological tests, be they ability tests, such as intelligence tests, or personality tests (see Dörner, et al., 1983, Stäudel, 1987). One possible explanation for these results is that microworlds demand other abilities than those assessed by standard tests. Another possibility highlighted by results from a study by Strohschneider (1986) is that the performance measures used to divide the subjects into extreme groups is not very reliable. If the performance variable is not reliable, we cannot, of course, expect very high correlations between this variables and tests.

Although lack of reliability may be part of the explanation for the poor correlations with standard tests, it is obviously not the whole story. There is some indication from studies using Moro, that a variable called *heuristic competence* might be related to performance in complex, dynamic systems. Heuristic competence can be seen as a general competence relating to complex systems. Stäudel (1987) developed a questionnaire for measuring this variable. It contained questions concerning people's habitual approach to problems, and required the subjects to answer questions concerning whether they would plan, whether they would check on the results of their actions, and so on.

Stäudel (1987) found significant correlations between the scores on her questionnaire and performance in Moro. Further support for the usefulness of this variable comes from a study by Schaub and Strohschneider (1989) which compared the performance of students and managers from German and Swiss companies. The managers performed better than the students, and the reason for this was not that they knew more about developing countries. Instead, the explanation could be found in differences in the manner in which they approached the task. Thus, the managers collected more information, planned better, and checked the results of their actions to a greater extent than the students. In short, the behavior of the managers suggested that they had better heuristic competence.

It is not clear how heuristic competence is acquired. When Dörner, Kreuzig, Reither and Stäudel (1983) tried to train their subjects in handling complex systems by giving them different kinds of short "courses" before attempting Lohhausen, they found no effects. Perhaps heuristic competence can be only be learned in context, i.e., as part of learning to work with actual systems. To understand how people acquire heuristic competence is obviously one of the important problems for future research with microworlds,

Some results from studies on expertise provide a clue. A study by Voss, Greene, Post and Penner (1983) on expertise in solving political science problems seems especially relevant here.

The task given to the subjects in the study was that of finding ways of improving the agricultural production in the Soviet Union. Voss et al. compared four groups of subjects: (1) subjects who were not only experts in political science but also experts in the particular domain relevant to the problems given in the experiment; (2) subjects who were experts in political science, but not in the particular domain studied in the experiment; and (3) a group of subjects who were experts in a totally different domain (chemistry). Finally, (4) there was a control group consisting of novices, i.e., subjects who were not experts in anything relevant to the purpose of the experiment.

The results showed that the chemistry experts performed at the same level as the novices. That is, expertise in problem solving in an area different from that of the problem did not help. The domain experts performed the best, followed by the experts in political science (but not on the Soviet Union). Voss et al concluded that expertise involves both content independent strategies and content dependent strategies. Domain experts had their knowledge hierarchically oriented in a manner so that the specific problems (such as lack of fertilizer) became subordinate to more abstract problems (such as lack of capital investment), and by solving these more abstract problems, the subjects could solve a host of specific problems, and did not have to consider all the specific problems one at a time.

These results suggest that heuristic competence is not totally separate from domain knowledge, or epistemic competence as Dörner, et al. called it. It seems necessary that the person has at least some experience in applying his heuristic competence in some context similar to that in which he is being tested. However, the exact relations remain obscure.

Demands Made by Dynamic Decision Tasks

Studies comparing the behavior of successful and less successful subjects in a given simulation have not yielded any surprising results: successful subjects collect more information, form better hypotheses and test them better, and check on the results of their decisions better than less successful subjects (e.g., Dörner, et al., 1983). In short, they show evidence of better heuristic competence. These are also behaviors that should lead to a better model of the system, and the results thus suggest that subjects who behave in a way that leads to a better model of the system will also succeed better in controlling it. This hypothesis also receives support from the finding that the successful subjects tended to ask more about causes and that they were much more directed at forming general concepts about the microworld. Less successful subjects seemed more oriented towards the next decision only, and less interested in forming a good model of the task (Dörner, et al., 1983). The successful subjects also seemed to be more prone to verbalization and subsequent self-reflection concerning their behavior. There seemed to be no differences in planning of decisions. The general conclusion from these results, then, seems to be that some subjects perform better because they are able to develop better *models* of the system that they are trying to control. From a control theory point of view, this is not surprising for, as one of the classical papers in that area states in its title, "A good regulator of a system must be a model of that system" (Conant & Ashby, 1970). Moreover, these results suggest that the better models are the result of superior heuristic competence, for the behaviors that characterize the better subjects are exactly those that are taken as evidence of this trait. However, the reason why only some of the subjects showed evidence of this is somewhat of a mystery.

The Pathologies of Decision Making

One of the most interesting results to come out of the analyses of individual differences has been the pathologies of decision making identified by Dörner (1980). These pathologies are behavior patterns exhibited by unsuccessful subjects in microworld experiments.

- The first of these pathologies is called *"thematic vagabonding "*, and it refers to a tendency to shift goals. Thus, a subject in, say, Moro, may start out doing something about the water supply, and when this fails to solve all the problems, he may then start working on education, and then shift back to working on the water supply, then move on to do something about storage facilities for the harvest, and so on. Thus, subjects who exhibit thematic vagabonding may shift from one part of the problem to another, all the time failing to work on the problem as a whole, which is what is required.

- The second pathology is called *"encystment"* and involves sticking to one specific goal that the subject feels competent to achieve. Thus, a subject who has great faith in education may stick to this method as the panacea for all problems in Moro.

- A third pathology is a general *refusal to make any decisions at all*, despite the need for decisions to reach the goals that have been set.

- The fourth pathology involves *blaming others for one's failures*. Many of these programs have simulated others who have to carry out the subject's commands. One of the most likely candidates here is, of course, the simulation itself, a response similar to that of blaming "the system" for one's failures.

- *Improper delegation*, i.e., delegating responsibility that cannot, or should not, be delegated, is a fifth pathology.

It is interesting to note that the five pathologies can be divided into two groups. The first group comprises the first two pathologies. These pathologies can be seen as failures of the goal formulation process. The second group comprises the last three pathologies, and the pathologies in this group may be seen as signs of a refusal to learn from experience. Thus, the first of the three pathologies in this group, refusal to make decisions, obviously means that the subject will not receive any new experience to learn on, while the last two involve rejecting whatever experience could have been used for learning.

According to Dörner (1989), these pathologies are not the causes of failures so much as the results of failure or rather of the low self confidence that results from failure. That is, they occur after the subjects have failed in gaining control, and they then, of course, further exacerbate the failure, leading to a worse and worse state of affairs.

These pathologies can be readily observed in various decision tasks. It would be important to monitor decision makers for these kinds of behavior since their occurrence might suggest that the problems facing the decision maker have become overwhelming (Brehmer, 1987).

THE EXPERIMENTAL APPROACH

As noted above, experiments with microworlds are concerned with the effects of *system characteristics* on the subject's behavior in the microworlds. Ideally, this undertaking requires a taxonomy of microworld characteristics. To find such a taxonomy is no simple task; systems theory offers a number of possibilities. Some attempts at defining taxonomies have been made (Brehmer, 1990, Funke, 1990), but here we will ignore these in favor of a simple list of system characteristics proposed by Brehmer and Allard (1990) which seems to cover most of the experiments made so far.

Brehmer and Allard (1990) proposed a list of six characteristics: complexity, feedback delays, feedback quality, relation between the characteristics of the process to be controlled and the control process, rate of change and possibilities for delegation of decision making power. We discuss each in turn.

- *Complexity*. Complexity is a difficult concept because it has no clear operationalization. Here, Brehmer and Allard (1990) followed Ashby's (1956) suggestion and defined the complexity of a system relative to the capacity of the mechanism that seeks to control that system. In the present case, the control mechanisms of interest are human beings, and we know that humans have limited capacity for processing information. Consequently, a preliminary definition of complexity could be cast in terms of the number of elements in the system to be controlled. However, in so doing, we need to distinguish between different kinds of elements. There are at least four different kinds of elements that need to be considered: goals, actions, side effects and processes to be controlled. Thus, complexity would be related to the number of processes to be controlled, the number of goals, the number of action alternatives, and the number of side effects. In the case of the fire fighting example above, complexity would thus vary with (1) the number of fires, (2) whether the subjects' goal is simply to put out the fires as quickly as possible, or whether there are other considerations as well, e.g., fires in some places may be more important than other fires, (3) the number of action alternatives, i.e., the number of fire fighting units (FFUs) and their characteristics, as well as (4) whether the agents used for putting out the fires had other effects as well, e.g., they might have negative effects on the flora and fauna of the forest.

- *Rate of change.* This refers to how quickly the process(es) to be controlled changes. Rate of changes may vary within very wide limits, with very fast tasks at the one end, such as that of performing a low level attack with a jet fighter, and extremely slow task at the other, such as that of controlling the economy of a country, with tasks such as that of fire fighting in between. An important question for research is, of course, whether systems with different rates of change are perceived and controlled in the same way.

- *Relation between the characteristics of the process to be controlled and the characteristics of the control processes.* This feature is unique to dynamic problems. As noted above, the problem facing the decision maker in dynamic decision making is to find a way of using one process to control another process. Figure 1 illustrates the nature of this problem for the fire fighting case.

  This particular combination of processes is, of course, only one of many possibilities. Each dynamic problem will have its own combination, and this characteristic will determine what the efficient strategy for the task is. For example, when the process to be controlled is a cancerous growth, and the means for control are some form of surgery, we are using a step function to control an exponential process.

- *Feedback delays.* The concept of delay refers to a slowing down, or lagging behind, in the transmission of energy, or information, in the system. This variable is, of course, quite complex in real dynamic systems, for delays may occur in many different places in the system. In the fire fighting example, the FFUs may be slow in responding to the commands (usually called *dead time* by control engineers), it will take time to execute the command (control engineers call this the *time constant* of the system), and the FFUs may be slow in reporting their actions to the fire chief. All of these different kinds of delay will show up in the same way to the decision maker: as a delay between the moment when a command has been given, and the moment when information about the results of this command is obtained.

  From the mere fact that there is such a difference, it is not possible to infer the location of the delay. For that, additional information about the system is needed. However, it is important know the source of the delay, for different kinds of delays require different kinds of models. Thus, to handle dead time and time constants the decision maker needs a model which makes it possible to predict when a decision is needed, for to compensate for these delays, the decision maker will have to make his or her decisions before there is any actual need for them. To compensate for slow reports, on the other, the decision maker needs a model which makes it possible to predict when the decisions have taken effect, and when his or her resources are freed for further decisions.

- *Quality of feedback information.* Information about the state of the system may vary in quality. Thus, it may be more or less complete because of the way in which the information system has been designed, or it may even be distorted, perhaps because the lower levels send incorrect reports about their activities.

- *Distribution of decision making capacity.* In a complex dynamic system, such as the fire fighting example, all decision making power may be centralized in the fire chief, or it may be distributed, so that the local FFU commanders are able to make some of the decisions.

  The possibility of delegating decision making power is of extreme importance when there are delays in the system. In that case, it is not possible to control the system centrally (unless one has a model that incorporates the delays), and delegation of decision making power to local units becomes mandatory.

Results

So far there is little research concerned with the effects of complexity. This may be surprising, since complexity is taken as an important defining characteristic of the systems of interest. However, upon reflection, it is obvious that there is something trivial about complexity as an independent variable. Since complexity cannot be defined independently of the capacity of the subject, experimental results become circular: either the subject cannot produce the requisite variety that the task demands, in which case the task is obviously complex, and the subject will perform badly, or, if he can produce the requisite variety needed, the subject will perform well, in which case the task is not complex.

Thus, it makes little sense to vary complexity experimentally. Instead, experimenters have focussed upon what subjects do in face of complexity. This research has already been reviewed above.

Feedback delays have been extensively investigated. The results show that feedback delays generally have detrimental effects on performance (Brehmer, 1990b). However, if subject is able to actually see the reasons for the delays, he or she may be able to compensate for the delays. For example, in Brehmer and Allard's fire fighting simulation, subjects are able to compensate for the time constants caused by the fact that it takes time for the fire fighting units to move and to extinguish fire (both which they can actually see form the screen) but not for dead time or for delays in reports, which they must infer (Brehmer, 1990b). The subjects do not compensate for such delays, even if they are informed that there are such delays (Brehmer & Allard, 1991). Moreover, they do not try to use delegation of decision making power to lower levels in the hierarchy as a means for combatting the effects of delays.

There are no systematic results on subjects' ability to cope with different forms of relations between control processes and processes to be controlled. It is clear that subjects can cope with the case of two linear processes, as is shown by the fact they are able to extinguish fires in the fire fighting simulation, but we know very little about other combinations.

Dörner (1980) has suggested that exponential processes are especially difficult to handle, but it is not clear whether the subjects have had to try to control such processes by means of linear processes (most man-made control processes seem either to work linearly or to involve step functions), or whether they would be difficult regardless of the nature of the control process. The fact that this particular aspect of dynamic tasks has not been studied experimentally means that we know very little about subjects' ability to obtain strategic knowledge in these tasks.

Research on dynamic decision making is obviously only in its infancy, and much remains to be done. Despite this, stable and interesting results have been obtained with respect to individual differences, pathologies of decision making and with respect to important system characteristics such as feedback delays.

CONCLUSIONS

The focus on process in European research on decision making has not only given us better insights into how decisions are made. Thus, this research shows that the hypothesis that human decision making is a constrained form of normative decision making, constrained because of incorrect probability estimates, inconsistent values and/or incorrect combination of probabilities and values, is not always correct. Instead, decision making may be viewed as a form of problem solving, where a person seeks a viable course of action.

The analogy between decision making and problem solving also points to new ways of understanding why human decisions are not always as good as they could be. It also shows what kinds of decision aids may be effective. The basic problem is that of knowledge. Thus, we must expect that decision makers will have little understanding of decision aids that just help them follow some normative procedure. Instead, they will want aids that increase their understanding of the decision problem. Consequently, we should expect that, for example, various forms of simulations that show the decision maker what the likely consequences of different courses of action would be, might be effective. Such decision aids are now being developed for process operators. As new simulation tools for personal computers are developed, such as STELLA (Richmond, Vescuso and Peterson, 1987) and the like, we must expect that simulation will soon be a more and more common decision aid.

REFERENCES

Ashby, W.R. 1956. *An introduction to cybernetics.* London: Chapman & Hall.

Ajzen, I. & Fishbein, M. 1980. *Understanding attitudes and predicting social behavior.* Englewood Cliffs, N. J.: Prentice-Hall.

Biel, A. & Montgomery, H. 1989. Scenario analysis and energy politics: The disclosure of causal structures in decision making. In Montgomery, H. & Svenson, 0. (Eds.) *Process and structure in human decision making.* Chichester: Wiley.

Borcherding, K. Larichev, 0. & Messick, D. (Eds.) 1990. *Contemporary issues in decision making.* Amsterdam: North-Holland.

Brehmer, B. 1988. Organization for decision making in complex systems. In L. Goodstein, H.B. Andersen, and S.E. Olesen (Eds.), *Tasks, errors and mental models.* London: Taylor & Francis.

Brehmer, A. 1989. Brännskadevård som ett hierarkiskt organiserat system. Uppsala University, Center for Human Computer Interaction. CMD Report No. 4.

Brehmer, B. 1990a. Towards a taxonomy for microworlds. In J. Rasmussen, B. Brehmer, M. de Montmollin & J. Leplat (Eds.), *Taxonomy for work domains. Proceedings of the first MOHAWC workshop.* Roskilde: Risö National Laboratory.

Brehmer, B. 1990b. Strategies in real-time, dynamic decision making. In R. Hogarth (Ed.), *Insights in decision making*. Chicago: University of Chicago Press.

Brehmer, B. & Allard, R. 1990. Dynamic decision making: Effects of feedback delays and complexity. In J. Rasmussen, B. Brehmer & J. Leplat (Eds.), *Distributed decision making: Cognitive models of cooperative work*. Chichester: Wiley.

Brehmer, B. & Allard, R. 1991. *Dead time and report delays in dynamic decision making*. Uppsala Psychological Reports.

Brehmer, B. & Dörner, D. (In press). Experiments with computer simulated microworlds: Escaping the narrow straits of the laboratory as well as the deep blue sea of the field study. *Computers in Human Behavior.*

Brehmer, B., Jungermann, H., Lourens, P. & Sevon, G. (Eds.) 1986. *New directions in research on decision making*. Amsterdam: North-Holland.

Conant, R.C. & Ashby, W.R. 1970. Every good regulator of a system must be a model of that system. *International Journal of System Science, 1,* 89-97.

Crozier, R. 1989. Postdecisional justification: The case of De Lorean. In Montgomery, H. & Svenson, 0. (Eds.) *Process and structure in human decision making*. Chichester: Wiley.

Dörner, D. 1980. On the problems people have in dealing with complexity. *Simulation and Games, 11,* 87-106.

Dörner, D. 1989 *Die Logik des Missl'ngens. Strategisches Denken in komplexen Situationen.* Reinbeck bei Hamburg: Rowohlt.

Dörner, D. 1991. *Über die Philosophie der Verwendung von Mikrowelten oder "Computerszenarios" in der pschologischen Forschung*. Working Paper No. 7. Berlin: Projektgruppe Kognitive Anthropologie. Max-Planck-Gesellschaft.

Dörner, D., Kreuzig, Reither, & Stäudel, T. 1983. *Lohhausen. Vom Umgang mit Unbestimmtheit und Komplexität.* Bern: Huber.

Dörner, D. Stäudel, T. & Strohschneider, S. 1986. *Moro-Programmdokumentation.* Memorandum 23. Lehrstuhl Psychologie II, Bamberg University.

Einhorn, H. J. & Hogarth, R. M. 1981. Behavioral decision theory: Processes of judgment and Choice. *Annual Review of Psychology, 32,* 53-88.

Ericsson, K. A. & Simon, H. A. 1984. *Protocol analysis: Verbal reports as data.* Cambridge, Mass.: MIT Press.

Funke, J. 1990. Systemmerkmale als Determinanten des Umgangs mit dynamischen Systemen. *Sprache & Kognition, 9,* 143-154.

Gallhofer, I. & Saris, W. 1989. Decision trees and decision rules in politics: The empirical decision analysis procedure. In Montgomery, H. & Svenson, 0. (Eds.), *Process and structure in human decision making*. Chichester: Wiley.

Huber, O. 1982. *Entscheiden als Problemlösen*. Bern: Huber.

Huber, O. 1989. Information processing operators in decision making. In Montgomery, H. & Svenson, 0. *(Eds.),Process and structure in human decision making*. Chichester: Wiley.

Kuhl, J. 1986. Human motivation: From decision making to action control. In Brehmer, B., Jungermann, H., Lourens, P. & Sevon, G. (Eds.), *New directions in research on decision making*. Amsterdam: North-Holland.

Montgomery, H. 1983. Decision rules and the search for a dominance structure: Towards a process model of decision making. In Humphreys, P., Svenson, 0. & Vari, A. (Eds.), *Analyzing and aiding decision processes*. Amsterdam: North-Holland.

Montgomery, H. 1989. From cognition to action: The search for dominance in decision making. In Montgomery, H. & Svenson, 0. (Eds.), *Process and structure in human decision making*. Chichester: Wiley.

Montgomery, H. & Svenson, 0. (Eds.) 1989. *Process and structure in human decision making*. Chichester: Wiley.

Newell, A. & Simon, H. 1972. *Human problem solving*. Englewood-Cliffs, N.J.: Prentice-Hall

Payne, J. W. 1976. Task complexity and contingent processing in decision making: An information search and protocol analysis. *Organizational Behavior and Human Performance, 16*, 366-87.

Rapoport, A. 1975. Research paradigms for the study of dynamic decision behavior. In D. Wendt & C. Vlek (Eds.), *Utility, probability and human decision making*. Dordrecht: Reidel.

Reichert, U. & Dörner, D. 1987. Heurismen beim Umgang mit einem "einfachen" dynamischen System. *Sprache & Kognition, 7*, 12-24.

Richmond, B., Vescuso, P. & Peterson, S. 1987. *An academic user's guide to STELLA*. Lyme, N.H.: High Performance Systems.

Rohrmann, B., Beach, L., Vlek, C. & Watson, S. (Eds.) 1989. *Advances in decision research*. Amsterdam: North-Holland.

Schaub, H. & Strohschneider, S. 1989. *Die Rolle heuristischen Wissens beim Umgang mit einem komplexen Problem oder: Können Manager wirklich besser managen*? Memorandum No. 71, Lehrstuhl Psychologie II, Bamberg University.

Sjöberg, L. 1980. Volitional problems in carrying through a difficult decision. *Acta Psychologica, 12*, 123-32.

Strohschneider, S. 1986. Zur Stabilität und Validität von Handeln in komplexen Realitätsbereichen. *Sprache & Kognition, 5,* 42,-48.

Stäudel, T. 1987. *Problemlösen, Emotionen und Kompetenz. Die Überprüfung eines integrative Konstrukts.* Regensburg: Roderer.

Svenmarck, P. & Brehmer, B. 1991. D^3 FIRE: An experimental paradigm for the study of distributed decision making. In B. Brehmer (Ed.), *Distributed decision making. Proceedings of the Third MOHAWC Workshop.* Risö National Laboratory. In press.

Svenson, O. 1989. Eliciting and analysing verbal protocols in process studies of judgment and decision making. In Montgomery, H. & Svenson, 0. (Eds.), *Process and structure in human decision making.* Chichester: Wiley.

Voss, J. F., Greene, T. R., Post, T. A. & Penner, B. C. 1983. Problem solving skill in the social sciences. In G. Bower (Ed.), *The psychology of learning and motivation. Vol. 17.* (Pp 165-213). San Diego: Academic Press.

Understanding and Aiding Military Decisions

by Martin A. Tolcott, formerly Head of the Psychological Sciences Division, Office of Naval Research, later Chief Scientist, Decision Science Consortium, Inc., now retired.

INTRODUCTION

The purpose of this paper is to provide an overview of the major trends in the study of decision making and the development of decision aids (or decision support systems) in the United States during the past approximately 35 years. It is not intended to be a comprehensive review of all such work. Rather, the most significant directions of psychological research effort will be described, focusing on work that has had clear implications for the design of aids to military decision making, and examples will be given of aids based upon the products of this research.

Although the automation of tasks can properly be regarded as a form of aiding, the emphasis in this paper is on situations in which the final decision is the responsibility of a human. The primary reason for this emphasis is that the major concern of psychologists is human behavior and how to exploit its strengths and overcome its weaknesses. But there is a practical reason as well. In military organizations clear responsibilities are assigned to humans at various levels and critical decisions, such as whether or not to engage the enemy and how to do so, will continue to be made by humans. This is the case even if a commander's decision is to activate an automatic target engagement system.

The two primary components of military decision making are situation assessment (what is happening) and action selection (what to do about it). Subsumed under these components are others, such as generating hypotheses to account for information being received, and generating and evaluating alternative actions. These steps are not always taken explicitly; often they are unnecessary because of previous training or experience, or because actions are specified by doctrine. Also, psychologists may describe these decision processes quite differently than military commanders. For the psychologist, the focus of attention has been to understand how inferences are made from uncertain evidence, that is, how one hypothesis about the situation comes to be regarded as more likely than others, and how actions are chosen for implementation, that is, why one action is considered better than other alternatives. Differing points of view between psychologists and military personnel, which will be described later, have led to serious shortcomings in early decision aiding efforts, but current trends and research approaches are reducing the divergence of attitudes and increasing the relevance of decision aiding techniques to a wider variety of military problems. This paper will describe the applications and shortcomings of decision aids, and what might be expected to come out of current and future efforts.

THE ERA OF RATIONAL MODELS

Normative Models of Rational Choice

During the period between about 1955 and 1975, decision aiding work was dominated by theories and models of so-called rational behavior. The idea was that there were correct ways to interpret evidence in assessing situations, and correct ways to assess the probable gains and losses of alternative courses of action, and that people could be helped to adopt these rational (or "prescriptive") procedures in order to improve their decisions. The procedures required the decision problem to be broken down (decomposed) into its elements, and judgments to be made about these smaller pieces of the problem. These judgments could then be recombined or aggregated by a formula (or a computer algorithm) in order to arrive at the best decision.

The most frequently used models of rational behavior were the following:

Bayesian inference models. Bayes' Theorem is a means of combining evidence from various sources to determine the likelihood of a hypothesis, much as a doctor would arrive at a diagnosis based on various symptoms and test results. The core of the Bayesian model is the equation:

$$P(H/D) = \frac{P(D/H) \times P(H)}{P(D)}$$

where P(H/D) is the posterior probability of the hypothesis, given the datum,

P(D/H) is the probability that the datum would occur if the hypothesis were true,

P(H) is the prior probability of the hypothesis, and

P(D) is the probability that the datum would occur regardless of the truth or falsity of the hypothesis.

The Bayesian model prescribes a way of incorporating new evidence into one's set of beliefs and revising one's assessment accordingly. Each time a P(H/D) is revised, it becomes the prior P(H) for the next revision. Ward Edwards introduced this model into psychology, and he and his followers were responsible for a rich period of research on various ways of eliciting and expressing the subjective judgments of probability that were needed when actual event frequencies were not available.

Subjective expected utility (SEU) theory. SEU theory provides a prescriptive model of rational choice, based initially on economic concepts of value, but modified to allow for individual differences in value judgments (subjective utilities). It also incorporates the element of uncertainty by taking into account the probabilities of uncertain events that would affect the outcomes. Thus, the subjective expected utility of an action equals the sum of the utilities of its anticipated outcomes multiplied by their probabilities of occurring, and the optimal choice among alternatives is that which results in the maximum SEU.

Multi-attribute utility (MAU) theory. MAU theory extends SEU to cases where the utility of an outcome depends on its value along several dimensions or criteria (or attributes). Thus, the utility of a certain make of car might depend on its score on such attributes as price, fuel economy, looks, handling qualities, etc., and people might differ as to the relative importance they place on each of these attributes. The MAU of an alternative is its aggregated utility across attributes, taking into account both the score on each attribute and the importance or weight assigned to each.

Aiding by Decomposition and Decision Analysis

Many of the early efforts to aid decision makers were based on rational models of the type described above. These efforts, which are still in use, involve decomposing the decision problem into its elements and modeling it in a form that makes explicit the choices open to the decision maker, the uncertainties that would affect the outcomes, and the outcomes themselves. This model usually takes the form of a decision tree. The decision maker is then asked to assign probabilities to the uncertain events (these might be actual frequencies based on past history, or they might be calculated by Bayesian analysis if appropriate, or simply estimated), and to evaluate or assign utilities to the various possible outcomes. The tree is then "folded back", that is, the appropriate arithmetic operations are performed to obtain the SEU of each possible choice, and the optimal choice becomes clear.

The decision tree method of analyzing decisions is not quite as simple as might be inferred from the above description. The reader is referred to Brown et al. (1974) for a detailed description of the method. The decision modeling itself is difficult; different decision makers as well as decision analysts might model the identical problem somewhat differently. When there are many uncertain future events the decision tree may become quite complex, and decision analysts have developed techniques for trimming unwieldy trees. The decomposed judgments about probabilities and utilities are not easily made by managers, most of whom are accustomed to making more global judgments. Nevertheless, the method has enjoyed some popularity; it is useful for certain types of decisions, and is still being used selectively in both military and industrial contexts.

In practice, most high-level decisions are not made by single individuals in isolation. More often the responsible individuals need inputs from others in their organizations, who may have special knowledge bearing on the problem, or who may be asked for their own judgments in an attempt to reach a consensus. Three techniques have been found useful by decision analysts to expedite their analyses when groups are involved:

Decision conferences. These are meetings of the key decision makers in the organization, led by the analyst, for the purpose of encouraging discussion relevant to the modeling and evaluation stages of the procedure. Several such conferences may be held over a period of time, allowing the analysis to develop gradually, and permitting the participants to review and revise their judgments, obtain new information if necessary, and continue the process until they are satisfied.

Computer-aided decision analysis. This is a procedure in which a template of the decision model is prepared in advance, and the analytic effort is devoted to revising it as necessary and filling in the detailed judgments (default values are sometimes used to avoid the need for difficult judgments). The model is fully developed by the analyst interacting with the computer, and if changes in estimated values are being contemplated by the participants, the effects of these changes can be rapidly computed (sensitivity analysis). Thus, the time of the participants can be spent more effectively in working to improve the precision of those judgments whose effects are most significant.

Group decision aids. In this technique, each participant is provided with a computer through which the individual's inputs can be transmitted to a central computer that displays the consensus. It has been claimed that this procedure avoids the potential domination of a group by aggressive or higher-ranking individuals, and therefore results in a truer consensus. However, this claim has never been validated.

THE ERA OF DESCRIPTIVE MODELS

Research on Biases in Judgment

From about 1965 to 1985, research psychologists accumulated much evidence that people do not normally make decisions in accordance with the rational models described above. It became a compelling research objective to determine how people actually make decisions. In this work, researchers took the rational models as a starting point, assumed that these models furnished the correct answers, and regarded deviations from the prescribed procedures as "heuristics" and deviations from the correct answers as "biases".

Perhaps the first to make explicit the fact of human deviation from so-called rationality in the context of decision making was Herbert Simon (1952). As early as 1957, Simon pointed out that in many situations people did not try to evaluate all their available action choices in order to maximize their SEU, but rather considered only as many alternatives as needed to discover one that satisfied them. He called this process "satisficing", and did not necessarily regard it as a heuristic leading to a biased result, but rather as a shortcut to meeting one's goals.

In 1968, Ward Edwards (1968) showed that people tended to be conservative in revising their judgments in the face of new evidence; they did not derive as much certainty from the new information as the Bayesian model specified. These early findings prompted a surge of research comparing human judgments with formal models and identifying a wide range of typical human decision biases. The work is extensively described in Kahneman et al. (1982). The following are a few examples of the findings reported:

Biases in probability estimation. The tendency toward conservatism has already been mentioned. In addition, people tend to underestimate the probability of very high frequency events and overestimate very low ones. They show a positive bias toward events that can be easily imagined or remembered ("availability" bias); they tend to ignore the base rate (the overall frequency of the event in some relevant reference population), and rely too heavily on small samples. They tend to ignore or undervalue new evidence that contradicts a previously-formed judgment ("confirmation" bias).

Biases in utility judgments. In assigning utilities to outcomes, people would normally be expected to show "transitivity" of judgment; that is, if A is preferred over B, and B over C, A should be preferred over C. Such is not always the case, however. There is great inconsistency in people's utility judgments, both within and between individuals. Judgments are influenced by many factors, including the way the problem is formulated or "framed". For example, mathematically identical choice problems will elicit different answers depending on whether the problem is formulated in terms of gains (e.g., number of lives saved) or losses (e.g., number of lives lost). People tend to be risk-aversive to gains and risk-seeking for losses.

Inadequate generation of alternatives. Whether the task is one of situation assessment or action selection, people have been shown to be very deficient in generating alternatives. In situation assessment they tend to generate only a few hypotheses based on early data, and find it difficult to enlarge their hypothesis set, even in the face of contradictory data. In generating alternative actions to solve a problem, they think of only a small fraction of the total number of possibilities.

Individual and task differences. People have been shown to differ in the way they approach decision problems; some tend to be analytical in their approach, breaking the problem into components and applying numerical weights and statistical procedures, while others tend to behave more intuitively, assessing the problem as a whole and coming to a solution very quickly. In multi-attribute utility analyses, some people prefer to consider one attribute at a time, judging its score across all the choices, while others prefer to consider one choice at a time, scoring it on all the attributes before considering other alternatives. Often they will eliminate an alternative if it fails to meet a threshold value on one of the attributes, ignoring the total utility. Different types of problems can also affect people's approaches; some problems tend to elicit a more analytic approach, while others (usually complex problems with many variables) elicit a more holistic approach.

Interactive and Adaptive Aiding to Reduce Biases

Personalized and prescriptive computer-based aids. The body of research on descriptive models led to an approach to decision aiding in which individual differences in style were recognized and accommodated, while at the same time the decision maker was guided away from biased solutions and procedures that might lead to them. In practice, allowing for individual differences took the form of a very flexible set of user-computer interface functions that permitted the user to employ a wide variety of approaches. For example (Cohen et al., 1987), in a MAU analysis the user could assign precise numerical values to the weights and scores of the attributes, or rank-order them, or adjust bar graphs to reflect approximate values, or even obtain a solution while ignoring some of the values. The user could also approach the problem attribute by attribute or alternative by alternative. This flexibility in accommodating to the user's preferred processing strategy is called "adaptive" or "personalized" aiding.

In order to guide the user away from procedures that could lead to incorrect solutions, the computer had a prescriptive program running in parallel with the user's procedures. This prescriptive aid kept track of the user's inputs or assumptions, warned when these assumptions could lead to trouble, and encouraged the user to seek and use additional information. For example, the computer might draw the user's attention to evidence that has not been considered. Or, a warning could be displayed if the user's conclusion was highly sensitive to a certain assumption and the user could be encouraged to view the problem under more than one assumption. A prompt might inform the user when an option has been eliminated because it falls short on one attribute if that option is significantly better on other attributes; similarly, the user could be informed when an option under consideration is dominated by other options (i.e., at least one other option is as good or better on each attribute). This aspect of the aid is called "prescriptive", and although it does not force the user to follow prescribed procedures, it provides guidance and warnings designed to reduce common biases.

Predictors and outcome calculators. In dynamic situations people are often unable to visualize how the situation will develop and, as noted earlier, are slow to change their views in the face of new evidence. For example, ignoring a target's recent maneuver will increase the error in estimates of its future location. Relatively simple computer algorithms can make these calculations and display on request the predicted future situation with the area of uncertainty clearly shown. Similarly, a computer aid can provide an estimate of the battle outcomes of alternative ground force tactics, given the current force ratios and deployments, and allow the user to make a more informed choice.

Artificial intelligence and expert systems. The field of artificial intelligence (AI) covers a broad range of efforts to design computers that can perform human-like activities. The major applications have been in artificial vision and other sensory processes, robotics, natural language processing, learning, and reasoning or problem solving. In the context of reasoning, AI generally takes the form of expert systems, which contain a set of rules for operating on data, and procedures or algorithms for triggering the appropriate rules when specific data are received. For the most part, AI systems are designed to replace the human. However, expert systems are designed as computer-based consultants to provide advice to users in a specific problem context, much as the personalized and prescriptive aid described above advises on decision procedures. The rules and logic by which AI and expert systems operate are usually obtained through extensive interviews with experts in the problem domain (subject matter experts), and the elicitation of this knowledge, free of common biases in judgment, is itself a difficult procedure which has only recently begun to receive significant research attention.

Graphic displays. Developments in computer and display technology during this period have allowed the use of vastly improved graphic and pictorial displays to aid decisions. In a military context such displays are usually geographic - battlefield terrain and force deployments, ship deployments and adjacent land areas, and aircraft routes to targets through varied terrains and threats. Rapid updating of the displays became possible, increasing their value in dynamically changing conditions. The use of color as a coding technique allowed distinctions to be made easily between enemy and friendly forces and among different types of units. Areas of coverage could be changed quickly, and cluttered displays could be uncluttered by selective filtering of unwanted information. Large screen displays could be used by teams located together, and rapid communications allowed distributed users to view identical displays at their separate workplaces and communicate about specific units through the displays themselves, reducing the likelihood of error. These developments were very important in increasing the ability of decision makers to think and talk about their decisions in their own terms.

CURRENT TRENDS

Decision Making in Natural Settings

Most of the experimental research conducted thus far had been performed in laboratory settings, usually in universities with students as experimental subjects, but sometimes in industrial or military laboratories with subjects who were trained in the tasks they were asked to perform. For the most part the laboratory tasks were significantly simplified versions of the actual decision tasks normally performed, in order to permit effective control of the variables of interest and precise measurement of the behavior of interest (such as estimating probabilities, assigning utilities, making inferences from evidence, etc.). Often the decision task itself was a contrived problem, deliberately designed to eliminate the effects of training and experience on the performance of the subjects, and to allow the experimenter to systematically vary certain factors.

Questions began to be raised about the applicability of the findings to real decision makers performing their jobs. In fact, such questions had begun to be addressed much earlier. For example, there had been an early interest by cognitive psychologists in how experts such as master chess players solved problems. Much of this early interest was related to efforts to develop automatic systems such as chess-playing machines, the idea being that computer programs should simulate the thinking of experts as much as possible.

In 1965, weather forecasters began to express their predictions of rain in probabilistic terms, and over the next ten years a mass of data on the accuracy of these predictions was accumulated. Murphy and Winkler (1977) published the results, and they showed remarkable accuracy. For example, when all the predictions of, say, 70% chance of rain were examined, it turned out that rain had actually occurred very close to 70% of the time. Forecast probabilities from 0 to 100% correlated almost perfectly with the observed relative frequencies. Weather forecasters were said to be very well "calibrated" as compared with population samples in most research.

Starting in the mid-1980's there began to emerge a strong research interest in studying decision making in natural settings that were more realistic than the laboratory settings of the previous era. The motivation for this work was partly an increasing scientific interest in understanding the cognitive processes of trained people doing their normal jobs, and partly the more applied goal of developing decision support systems that were more compatible with their users' actual decision processes. This movement has grown to the point where it is perhaps the dominant theme in current research activity. The theme is characterized by several features.

Increasing access to trained subjects. It is always difficult to obtain access to people in their working environment for research purposes, especially when the researcher wants to explore thinking processes and must often interrupt the work to ask questions. However, a growing interest in work on cognition and decision support systems has opened opportunities to observe people either doing their jobs or in their training environment conducting realistic exercises.

For example, Klein and his colleagues have observed urban and wildland firefighting teams in their work situation (Klein et al., 1986; Taynor et al., 1987), as well as Army personnel engaged in battlefield planning exercises (Thordsen et al., 1987). Tolcott et al. (1988, 1989a, 1989b) have worked with Army intelligence analysts at the Army Intelligence Center and School, asking them to solve simulated problems involving the predicted location of an enemy attack. Observations have been made aboard Navy ships and at Navy training facilities for Combat Information Center (CIC) teams, as well as at Air Force tactical planning centers. Some of this work has been conducted at an unclassified level and findings have been published in the open literature.

Focus on the total decision and underlying cognitive processes. Instead of breaking down the decision task into smaller components, the current trend, even in simulated exercises, is to present the task as realistically as possible, and by analysis of verbal protocols, interview data, or recorded communications, try to understand the underlying decision processes. This work is shedding new light on how experienced personnel solve their decision problems.

Recent findings. The results of this work have added to our theoretical knowledge and are forming a basis for new approaches to decision aiding. In Klein's work, for example, it was found that trained personnel normally do not consider and evaluate several alternative actions. Rather, they recognize certain aspects of a situation that are similar to situations they have dealt with before, and they consider the action that was successful on the previous occasion; they evaluate that action for its applicability to the present situation, through mental simulation of the acts and outcomes, and if they conclude that it will work, they proceed. If not, they modify it to fit the present circumstances, and re-evaluate it. This model, called "recognition-primed decision making", is similar in many respects to Simon's early concept of "satisficing", except that it incorporates the notion of recognition of previously met situations.

In my own work with intelligence analysts (Tolcott et al., 1989b), we were interested in the effects of early judgments on the processing of later information in an evolving situation. We found that subjects differed in their early judgments of where the enemy was going to attack, but regardless of their first estimate their confidence in it tended to rise with subsequent information even though they all received the same information. They showed an overwhelming tendency to regard the new information as confirming their early judgments (confirmation bias). 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 et al., 1988) it was found that the confirmation bias could be reduced by a brief orientation on biases in judgment, and by displays that made explicit the uncertainties about enemy unit locations. It was also found (Tolcott et al., 1989a) that if the subjects were allowed to select the items they wanted information about, they were more likely to pay attention to disconfirming evidence than when they were the passive recipients of information. These findings have implications for training and operational procedures as well as for decision-aiding displays.

Decision Making Under Stress

In recent years the U.S. Navy has experienced two incidents in the Persian Gulf that have increased official interest in decision research, particularly under stressful conditions. In the first incident the U.S.S. Stark did not defend itself against an incoming target, and suffered serious casualties. In the second, the U.S.S. Vincennes did choose to fire against an incoming target, and destroyed a civilian airliner. These incidents illustrate two sides of the problem of dealing with ambiguous information under stress, and they were partially responsible for stimulating a new Navy program on Tactical Decision Making Under Stress (TADMUS) [*cf.* the article in this issue of ESNIB by Cannon-Bowers, et al.].

Although there has been substantial research on the physiological effects of stress, it has been difficult to simulate the kinds of stress thought to characterize the combat environment; therefore we know little about the effects of stress on cognitive performance in combat. In the TADMUS program there will be no attempt to simulate conditions that might produce fear. Rather, the approach is to simulate conditions that appear to be typical in Combat Information Centers conducting air defense operations. These are the following:

Short decision time. Especially in constricted waters like the Persian Gulf, where threats can be launched from nearby land sites, the time for decision making is extremely short - a matter of minutes or even seconds. The lack of time to carefully sift and evaluate evidence is a major source of stress affecting decision makers. The TADMUS program will focus on decision making with short time limits.

High workload. During the Vincennes incident the Captain and the CIC team were busy conducting combat operations against small surface vessels when the approaching aircraft was detected. The high workload contributed to the stress of the situation by limiting the amount of attention that could be paid to the information being gathered about the approaching aircraft. Again, the TADMUS program will be concerned with decision making under high workload conditions.

Ambiguous and incomplete information. As described earlier, there is a great deal of research on decision making where the information provided is incomplete and ambiguous. What distinguishes the TADMUS program is that this feature will be combined with the characteristics of short decision time and high workload.

Other stressors. Auditory overload, noise, sustained operations, and sleep loss will also be used as stressors.

Group Decision Making

A previous Navy program examined group decision making under conditions where the decision responsibilities were distributed among team members in different locations. The work produced interesting findings about how the group communicates and reorganizes its responsibilities under different conditions. New models have been developed to describe the group's information transmission characteristics; a Petri net model has been found to have some useful properties for this purpose.

The TADMUS program will build upon these findings, and extend the focus to hierarchical groups, in which team members feed specialized pieces of information up to a higher level officer who must evaluate and synthesize the information to assess the total situation.

Integration of Decision Aiding and Training

As was shown in the studies of weather forecasters and intelligence analysts, there is some evidence that experience and training can improve decision making. It is not always clear whether there is more to be gained by developing decision support systems or by focusing on training to raise the level of performance. There could be interesting tradeoffs between the two approaches. For example, general purpose decision aids have not been very successful; thus, a decision aid might improve performance significantly, but only in a narrowly defined task. Training of course can be narrowly focused as well. However, training in broad concepts such as inference, the handling of uncertainty, the ability to respond rapidly in unanticipated situations, etc., might result in less improvement than decision aids, but could be beneficial over a much broader set of problems. Most of the research findings suggest that training in reasoning skills does not transfer to a wide variety of situations. But within the constrained context of tactical decision making, there might be performance improvements across many tasks. In any event, the TADMUS program offers an opportunity to examine these tradeoffs, since it will investigate training and decision aiding in the same program.

MILITARY APPLICATIONS

Aiding Based on Rational Models

Earlier in this paper the use of rational models such as Bayes' Theorem and subjective expected utility theory was described in general terms. The following are some examples of the application of these models to military decisions. (References are provided where descriptive publications are available for general distribution).

System evaluation. When a new system is in its early development stage there are often several alternative versions that are possible, each with its advantages and disadvantages. To oversimplify an actual case, the requirements for a field army radio included specifications for minimum acceptable transmission range and maximum weight. Early design studies suggested that greatly increased range would be possible, but at the cost of added weight which would reduce the unit's portability in the field. The issue was the utility of the added range and the dis-utility of the added weight. Since the radio would be used in a variety of missions, and the importance of portability as well as range varied with the mission, it was necessary to take into account the frequency or expected probability of occurrence of the various missions. A decision conference was held, the problem was decomposed and modeled, and the necessary judgments were obtained to allow the optimum design to be specified.

Investment strategy. A similar approach was used to decide how to allocate money among a large group of military development projects. The relative utility of each proposed system to each mission was determined, and the missions were rated in terms of overall criticality, an attribute that combined frequency and importance. Each proposed system was also judged as to the probability that its development would be successful, thus taking into account the developmental risk, and of course the costs were considered. The results of the analysis provided a prioritized list of projects that formed the basis for the final resource allocation decision.

Evaluation of tactics. The Army has developed a decision aid called TACVAL (for TACtical eVALuation) to help a tactical commander evaluate alternative courses of action (COA's) on the basis of a multi-attribute utility analysis. TACVAL provides the user with a MAU structure, and the user completes the model by entering the options, attribute weights and scores. TACVAL then computes and displays a numerical ranking for each COA. Although such an aid would rarely if ever be used during combat, an aid similar to TACVAL has been used at the Army War College as a training tool, enhancing students' understanding of the implicit judgmental processes involved in tactics planning.

Target classification. Systems to aid in determining if targets are hostile or friendly have incorporated Bayesian models to combine evidence from various sources into probability estimates, in the context of both Navy anti-submarine warfare and Army air defense. In air defense the problem is complicated by the presence of a large number of friendly aircraft in the area, presenting the risk of fratricide. A recently proposed feature (Ralston, 1989) is one that allows a combat commander to set a threshold for firing, depending on his decision as to an acceptable ratio between the number of hostile and friendly aircraft likely to be killed; the computation takes into account the performance effectiveness of the various sensors as well as the relative numbers of hostiles and friendlies in the area.

Intelligence analysis. The Army has developed an experimental Bayesian aid for intelligence analysis called BAUDI, (for Bayesian Aid for UpDating Intelligence information). The user specifies an initial set of hypotheses and the initial probabilities of each; as each new item of information is received the user enters the probability of receiving the information item, given each hypothesis. Based on these inputs, BAUDI calculates the posterior probabilities for each hypothesis. The experimental work (Adelman et al., 1982) showed that intelligence analysts using the aid were more consistent with normative Bayesian results, and were better able to distinguish the relative likelihoods of alternative hypotheses, than analysts not using the aid.

Limitations. As seen by the examples described above, aids based on rational models can be effective in certain military situations. However, they have shortcomings that limit their applicability.

There are two major limitations of the SEU-based decision analysis technique. The first is the relatively long time required to structure or model the decision problem and obtain the judgments needed to derive a solution. The technique can be used effectively in long-range planning problems which do not require rapid solutions and the critical features of which are not likely to change, in other words, where the judgments obtained will remain fairly stable over time. However, it is not well suited to handle tactical decisions in a fast-paced, rapidly changing situation, where event probabilities and even utility judgments may change as the battlefield situation changes.

The second problem is that the judgments required are unnatural and difficult for many people. This is not to say that battlefield commanders never estimate the likelihood of events or judge the utility of various outcomes. But they tend to make these judgments implicitly, without assigning numerical values, and on a more intuitive basis than the decision analytic procedure calls for. However, as we have noted, these models can play a useful role as training tools.

The Bayesian inference model, on the other hand, has been used effectively in several systems, primarily those concerned with target classification. However, even here there are limitations. The most serious is that when the situation changes (e.g., a sensor becomes less reliable, or the enemy introduces new units or tactics), the original model may no longer be appropriate. It may be quite difficult for the user to change the model, and a system's ability to modify itself adds significantly to its complexity.

However, as we shall see below, rational models can be combined with AI features that furnish a basis for changing the models, and in fact can make these changes automatically, thus providing a more *rapidly adaptable aid that is more useful* in dynamically changing situations.

Aids Based on Descriptive Models

A more recent approach to decision aiding is represented by systems that allow the users to solve problems in their own way, but include features that reduce the likelihood of errors. In these systems, when user judgments are required, they can be made in familiar military terms, making the systems more acceptable.

Some examples of these recently developed aids are the following.

Army battlefield planning. As described earlier, it has been found that intelligence analysts, when predicting the location of an enemy attack, tend to ignore evidence that contradicts their early judgments. One approach to reducing this tendency is found in a prototype aid for estimating enemy courses of action (AI/ENCOA) that incorporates both AI and MAU features. The MAU model is a hierarchical set of attributes for evaluating the utility of alternative enemy options. However, the aid also incorporates two other features: one is a user interface that queries users about the current situation and allows them to respond in military terms, with which they are familiar; the other is a set of rules that translate the situation descriptions into appropriate weights and scores in the MAU model. Thus, updating the model is made easier.

A somewhat similar approach is used in an aid to support Army personnel in formulating concepts of operations. Called TACPLAN (for TACtical PLANner), the system permits planners to enter their guidance and mission definition, and elicits judgments about military aspects of the situation (e.g., terrain characteristics, relative combat power, etc.). TACPLAN matches these judgments against criteria defined in various rule bases. A video disc-based interface allows the user to draw various candidate courses of action directly onto a map display, and TACPLAN, after checking the rule bases, reports back to the planner about problems and opportunities presented by the candidate courses of action.

Air Force mission and route planning. The Air Force has developed a set of prototype aids to be used in tactical planning. The most comprehensive is TEMPLAR (Tactical Expert Mission PLAnneR), which helps in selecting targets, selecting weapons, and assigning forces. It also accomplishes the routine and tedious task of preparing the Air Tasking Order (ATO), the detailed schedule of sorties distributed to various air bases. The concept was that TEMPLAR could receive the outputs of other, more specialized, aids, or it could generate the required data either automatically or with manual inputs.

One of TEMPLAR's supporting aids is ESCMA (Enemy Sortie Capability Measurement Aid), which uses an expert knowledge base to identify key enemy air base targets by determining their sortie generation rate, according to their location, types of aircraft, weather, etc. Another targeting aid is SPEA (See and Project Enemy Activity), which predicts future enemy ground troop positions based on conditions affecting their movement (weather, terrain, equipment), using an expert knowledge base of rules about the effects of these factors and their interactions.

In its final version, TEMPLAR is to combine this targeting information with data about own force status, and select the number and types of aircraft for the next day's strikes, their weapons, the air bases from which they should be launched and their targets. The planner could override or modify any portion of the plan, and could try out proposed modifications to determine their effects on the overall plan before adopting them.

Another Air Force aid is the Route Planning Aid (RPA), which uses information about terrain, threat location and threat type to select the best strike routes and altitudes for minimum lethality. It would also evaluate routes proposed by the user. Proposed improvements to RPA would allow use of other criteria such as time and fuel consumption, at the request of the user.

Navy command and control. The Navy's approach to aiding decisions in command and control (C2) systems has been, in general, to rely on computational aids that perform complex and burdensome calculations, reducing the workload on personnel, speeding up the dissemination and display of information, and providing more time for command decision making. Illustrative of the Navy's approach is the Joint Operational Tactical System (JOTS) (1990). JOTS receives tactical data from many sources, distills the data into a local database, displays it in various ways, and distributes it to other systems. A selection of Tactical Decision Aids (TDA's) is available for use with the system; the TDA's process the data to provide a variety of computations, predictions, and conversions of the data into different forms. For example, one of these aids (QUICK CPA) calculates and plots the closest point of approach between two tracks selected by the user. Another, QUICK INTERCEPT, calculates and plots the intercept solution between two tracks. A somewhat more complex computation is performed by IREPS (Integrated Refraction Effects Prediction System), which predicts radar ranges as a function of the environmental conditions that can affect them. In all three cases described here the results of the calculations can be displayed in tabular as well as pictorial or graphic form.

A more sophisticated Navy decision aid is one which determines the best sonobuoy pattern for an anti-submarine warfare (ASW) airplane to drop when it is trying to localize the position of a submarine target. The aid takes into account the initial detection data, the submarine's most likely area, and any search constraints entered by the user, and it computes the best sonobuoy pattern as well as the navigational information needed by the aircraft to execute the plan. Users can also enter their own plans and have the aid evaluate them. A strong feature of this system is that the user can specify the criteria for evaluation; the aid can use probability of detection, estimated mean time to first detection, mean time the contact will be held, or aircraft deployment time, or it can combine these, with user-determined weights, into an overall measure of probability of success.

Limitations. The advantages of these recently developed decision aids over earlier ones are that the inputs required of users and the outputs furnished are in operational terms familiar to the users, and that they provide a good deal of flexibility, allowing for differences among users as well as among situations. These features are consistent with the findings of research on descriptive models of decision making.

However, even these aids are limited in their applicability, to combat planning or to training rather than to fast-paced tactical operations. One reason is that the flexibility that allows for user choice demands user inputs, which take time and often result in frustration. This may be tolerable during planning because there is sufficient time, but during combat it becomes unacceptable. Furthermore, in order for the knowledge (rules, logic) of a decision aid to be accommodated in a reasonably sized data base, that knowledge domain must be severely circumscribed. Thus, even the recent decision aids that can adapt to changing conditions operate within a very limited problem domain. In this connection the TEMPLAR concept of a high-level aid supported by lower-level specialty aids may be the most promising direction for future development. The aids that are most likely to enjoy widespread use during combat are the computational aids that are part of the Navy JOTS system.

The Potential of Aids Based on Current Research

Current research on decision making in naturalistic settings by trained personnel could lead to significant improvements in the aiding of military decisions. Some of these are described below.

Compatibility with cognitive processes. It is now generally acknowledged that trained personnel do not go through a lengthy analytical process to find the optimum solution each time they are faced with a decision. A frequently used shortcut is to recognize similarities between the current situation and ones they have faced in the past, to mentally select a course of action that was satisfactory on previous occasions, and to take that action if it seems satisfactory at present or modify or replace it if it does not. Thus far these findings have not led to new concepts for decision aiding, but one can imagine a system in which a variety of typical conditions are stored in memory, and the anticipated pattern of events under each condition can be displayed. The current situation can then be matched against the stored templates to aid the user in finding the most appropriate model. An additional feature might highlight the differences between the current situation and the closest one in memory, allowing the user to judge how important these differences are, and reject or modify the implied action accordingly. A more sophisticated addition might be a display of the action or actions that have been successful in the past, with perhaps indicators of what modifications might be necessary to compensate for differences. Although the storage requirements of such a system would be large, the obvious advantage would be that the memory would not be limited to that of the individual user but could encompass the experiences of many experts. Indeed, the concept is similar to that of modern expert systems, but the user inputs and system outputs would be more consistent with the cognitive processes underlying recognition primed reasoning.

Applicability to fast-paced, unstructured situations. The applicability of the concept described above would still be limited to combat situations which could be reasonably well predicted in advance. Human judgment will be required to handle unanticipated problems in fast-paced combat. Perhaps the best approach to aiding under these circumstances is the Navy's approach, that is, to increasingly automate the tedious computations and predictions that consume scarce time, and thereby free the humans to make the necessary judgments. When a new situation tends to recur with minor variations, one can envision a system that learns over time, and evolves into the type of expert recognition-primed decision aid described above. However, to the extent that a situation is unique, human judgment will be needed.

Training for decision making. Since the requirement for human judgment will never be eliminated, it seems clear that training of decision makers will always be needed. However, the training must focus not only on the ability to recognize situations that have been faced before (although this will continue to be necessary), but on the cognitive processes that will be needed to handle new situations quickly. The training program should present a wide variety of scenarios, and require decisions to be made under conditions of short time and high workload. Students should be taught to identify the most important cues to situation assessment, the largest areas of uncertainty, the best sources of information to reduce the uncertainty, and the risks involved in actions that seem satisfactory. They should also be made aware of common biases in judgment, and learn how these biases may show themselves in the various scenarios being presented. Current training emphasizes knowledge of force capabilities, tactical doctrine, and procedures (such as equipment operation, terrain analysis, communications, report preparation, etc.) What is suggested here is that some attention be devoted to training in the reasoning skills essential to good command performance.

FUTURE REQUIREMENTS

Techniques for Evaluating Decision Effectiveness

A major requirement is for an improved technology to evaluate the effectiveness of decision making. Such a technology is essential in order to evaluate the effects of decision aids and training.

In the real world, decisions are usually judged by whether the outcome was successful. This is risky for two reasons: (1) in an uncertain world a good decision may lead to a bad outcome and vice versa, and the same situation is unlikely to occur frequently enough to determine if the decision leads to a good outcome more often; (2) many decisions, especially in combat, result in unintended consequences (e.g., the Kurdish refugee problem in Iraq), which complicate the assessment of outcome utility especially if the unintended consequences take a long time to emerge.

In laboratory research, the rational models are usually taken as the criteria for decision effectiveness, despite the fact that the problem solver may not have been looking for the optimum solution. Often, adherence to the so-called rational process is used as the measure, even though decision makers can often arrive at the best solution through short-cuts that they may not be able to describe. "Thinking aloud" and protocol analysis are useful, although cumbersome, procedures for descriptive studies, but less useful for evaluation.

In naturalistic studies it is often possible to compare the observed decisions with those arrived at by a panel of experts. Pre-stored "school" solutions are useful for problems that are used frequently, but obtaining experts' solutions to a wide variety of one-of-a-kind scenarios could be a cumbersome process, although perhaps necessary.

A standardized technique for evaluating decision effectiveness may in fact not be attainable, but it should be possible to agree on a set of criteria that could apply even to an expert or school solution. The explication of such criteria would provide clearer goals for the future development of both decision aids and training.

Training in the Use of Decision Aids

Another area of work needing attention is that of training decision makers in how to use decision aids effectively. To the extent that computational aids are embedded in major units of equipment, the necessary training is incorporated into the training in equipment operation. However, many of the more sophisticated aids are likely to be designed as stand-alone systems, and special training will have to be provided to make sure the user is aware of the various ways in which the aid can be used - the opportunities it offers - as well as its limitations.

One future trend in decision aid development will almost certainly be in the direction of aids that learn, or adapt, or in some way modify themselves as the situation demands. Users must have confidence in their equipment, and users of decision aids lose confidence if they do not understand the rules or logic by which the aids operate. This problem will become worse as the aids themselves become adaptable and possibly even change their operating logic. Expert systems designed as tutors usually give the user a chance to request an explanation of the aid's advice, often at several different levels of detail; these explanations are embedded in the system. It will become increasingly important to find ways to embed in decision aids the appropriate explanations to users. Furthermore, in the case of self-learning or adaptable aids, the users should receive warnings or advisories whenever there is a change in the system's logic.

Thus, the challenge here is to develop not only initial training strategies for people who will be using the aid, but to embed into the aid the continuing refresher training needed to maintain user skills as well as to update the explanations of the aid's logic as often as it becomes modified. A decision aid that is not well understood will not be used.

REFERENCES

Adelman, L., Donnell, M.L., Phelps, R.H. & Patterson, J.F. 1982. An iterative Bayesian aid: Toward improving the user-aid and user-organization interfaces. *IEEE Transactions on Systems, Man and Cybernetics,* November/December.

Brown, R.V., Kahr, A.S. & Peterson, C. 1974. *Decision Analysis for the Manager.* New York: Holt, Rinehart and Winston, Inc.

Cohen, M., Laskey, K.B. & Tolcott, M.A. 1987. *A personalized and prescriptive decision aid for choice from a database of options.* Reston, VA: Decision Science Consortium, Inc., Technical Report 87-18.

Edwards, W. 1968. Conservatism in human information processing. In B. Kleinmuntz (Ed.) *Formal* *representation of human judgment.* New York: Wiley, 17-52.

JOTS II 1.0 User's Guide. 1990. INRI Inter-National Research Institute, Yorktown, Virginia.

Kahneman, D., Slovic, P. & Tversky, A. (Eds.) 1982. *Judgment under uncertainty: Heuristics and biases.* Cambridge: Cambridge University Press.

Klein, G.A., Calderwood, R. & Clinton-Cirocco, A. 1986. Rapid decision making on the fire ground. *Proceedings of the Human Factors Society 30th Annual Meeting.* Dayton, Ohio: Human Factors Society.

Murphy, A. H. & Winkler, R.L. 1977. Can weather forecasters formulate reliable probability forecasts of precipitation and temperature? *National Weather Digest, 2,* 2-9.

Ralston, J.M. 1989. Objective rules of engagement and sensor fusion in combat identification systems. *Proceedings of the 1989 Symposium on Command and Control Research.* McLean, VA: Science Applications International Corp.

Simon, H. 1952. *Models of Man: Social and Rational.* New York: Wiley.

Taynor, J., Klein, G.A. & Thordsen, M. 1987. *Distributed decision makinq in wildland firefightinq.* KATR-858(A)-87-04F Yellow Springs, Ohio: Klein Associates, Inc.

Thordsen, M., Galushka, J., Klein, G.A., Young, S. & Brezovic, C.P. 1987. *Distributed decision making in a command and control planning environment.* KATR-863(C)-87-08F Yellow Springs, Ohio: Klein Associates, Inc.

Tolcott, M.A. & Marvin, F.F. 1988. *Reducing the confirmation bias in an evolving situation.* Interim Technical Report 88-11, Reston, VA: Decision Science Consortium, Inc.

Tolcott, M.A., Marvin, F.F. & Bresnick, T.A. 1989a. The confirmation bias in evolving situations. Proceedings of the 1989 Symposium on Command and Control Research. McLean, VA: Science Applications International Corp.

Tolcott, M.A., Marvin, F.F. & Lehner, P.E. 1989b. Expert decision making in evolving situations. *IEEE Transactions on Systems, Man and Cybernetics, 19* (3).

Improving Tactical Decision Making Under Stress: Research Directions and Applied Implications

by Janis A. Cannon-Bowers and Eduardo Salas, Naval Training Systems Center, Orlando, FL and Jeffrey D. Grossman, Naval Command, Control and Ocean Surveillance Center, San Diego, CA.

ABSTRACT

Tactical decision making teams in the modern warfare environment are faced with situations characterized by rapidly unfolding events, multiple plausible hypotheses, high information ambiguity, extreme time pressure, and severe (often catastrophic) consequences for errors. The Tactical Decision Making Under Stress (TADMUS) project was initiated to address the problem of maintaining individual and team decision making in such an environment by applying recent advances in decision theory, display design, and training. The purpose of this paper is to describe the objectives and approach of the TADMUS project, and to highlight new and emerging technologies that are being used as a basis for the research.

The TADMUS program was designed to: define the decision problem facing Navy tactical teams, and develop measures of tactical decision making performance; collect empirical data to document the impact of stress on decision making; and develop principles for decision support, information display, training system design and simulation that will mitigate these stress effects. Several emerging areas of research are being exploited to accomplish these objectives, including: recognition primed decision theory, shared mental models, human performance modeling, and team training.

INTRODUCTION

Demands on the human decision maker in military tactical environments are becoming more complicated with advances in technology and changes in the world order. Modern combat scenarios are often characterized by rapidly evolving and changing conditions, severe time compression, and high degrees of ambiguity and uncertainty. In addition, such situations often present the decision maker with an overwhelming amount of data, and require the coordinated performance of a team of operators who must gather, process, integrate, communicate and act upon these data in support of a tactical decision. A variety of other stressors (both environmental and psychological) also exist in the operational setting, not the least of which is the potentially catastrophic cost of making an error. High levels of these operational stressors can severely degrade individual and team performance. Coupled with the fact that the modern military scenario is likely to be complex and multi-national, these factors have provided unprecedented demands on the human decision maker.

An example may help to illustrate this contention. One of the tasks facing a team of operators in a Navy combat information center (CIC) is to defend the ship against hostile aircraft and missiles. This task is accomplished by a hierarchically-structured team of operators/decision makers, with final decision authority retained by a single individual. Team members perform a variety of functions in support of the final decision: they operate sensor consoles to detect aircraft, integrate information collected regarding the aircraft's intent, make decisions about how and when to seek additional information, and make decisions about how and when to transmit pertinent situation assessment information. Once information is passed to a final decision maker, it must be considered against the situational constraints (e.g., rules of engagement) and potential consequences before a decision can be reached or action taken.

In order to function in such a situation, team members must understand the system at several levels. First, they must understand the dynamics and control of the equipment (both hardware and software) with which they are interacting to extract information. Second, they must understand the demands of the task and how to accomplish it (e.g., the significance of information, types of information required, strategies to combine information, necessary procedures, etc.). They must also understand how various facets of the environment affect the task and task demands; as for example, when workload increases as a function of air traffic in the area, or when radar reception is affected by weather conditions. Third, they must understand their role in the task, that is, what their particular contribution is, how they must interact with other team members, who requires particular classes of information, and so forth.

Finally, in order to perform optimally, tactical team members must be familiar with the knowledge, skills, abilities, preferences and other task-relevant attributes of their teammates. This is due to the fact that expectations for the behavior of their teammates will vary as a function of the individuals who comprise the team. When working with a particularly competent teammate, for example, a team member may alter his behavior so that it is consistent with how he thinks that teammate will perform.

It should be obvious at this point that the modern tactical decision making environment is a complicated one that places extraordinary demands on human operators. Unfortunately, past research into decision making and related fields has not yielded results that are readily applicable to improving performance in the modern combat scenario. Recently, a program of research was initiated to address this critical need. Called the Tactical Decision Making Under Stress (TADMUS) project, this effort seeks to apply recent advances in decision theory, individual and team training, and information display to the problem of enhancing the quality of tactical decision making under conditions of stress.

The purpose of this paper is to first describe briefly the TADMUS program, and then to focus in more detail on research directions that have emerged as a function of the TADMUS effort to date. These research directions include: gaining an understanding of decision making and stress in the operational environment, measuring tactical decision making performance, specifying training and simulation principles for tactical decision making, and developing decision support and display principles for tactical decision making.

THE TADMUS PROGRAM

The TADMUS program, which is sponsored by the U.S. Office of Naval Technology, is being accomplished through a cooperative program in human factors and training involving two principal Navy laboratories (the Naval Training Systems Center and the Naval Command, Control and Ocean Surveillance Center) as well as other Navy, industrial and academic organizations. The technology developed under this program will be demonstrated and evaluated in the context of Navy anti-air warfare scenarios. From this, general principles will be developed that will be applicable to other domains. TADMUS is comprised of five interrelated thrusts:

1. *Definition and Measurement of Critical Decision Tasks.* In order to evaluate properly any new technology being developed, it is important to provide the context of a realistic operational scenario. This task involves: defining the operational tasks; establishing laboratories in which to study those tasks; developing performance measurement capability; and developing models of the decision making processes for that operational environment.
2. *Examination of Stress Effects on Decision Making.* The objective of this task is to understand how combat-related stress affects tactical decision making. Prior to the evaluation of the technology under development in this program, a baseline is needed that describes how performance degrades under various stressors. Work in this task will involve selecting a number of stressors for investigation, determining which stressors should be used as approximations to actual combat stress, and determining how to quantify their effects.
3. *Development of Training and Simulation Principles.* The focus of this task will be on developing and demonstrating a variety of individual and team training strategies and techniques to minimize the adverse effects of stress. Products of this task will include principles for: overtraining decision making skills; training decision makers in pattern recognition; training interventions that will attenuate the effects of stress on team performance; training leadership skills.
4. *Development of Decision Support Principles.* An experimental decision support system will be produced which incorporates sufficient flexibility to permit extensive exploration of alternative concepts and architectures. The prototype will be evaluated in simulated tactical environments, initially in laboratory settings and later during at-sea exercises. Additional products of this task will be general principles for advanced decision support systems to enhance human performance under stress.
5. *Development of Display Principles.* Work will be undertaken to examine man-machine interface concepts that maximize the effectiveness of tactical decision aids under stressful conditions. Products of this task will include display principles for: predictive displays; situation assessment; option generation; resolution of conflicting or ambiguous information; cognitive consistency among team members.

RESEARCH DIRECTIONS

A number of research topics have emerged from the five thrusts described above with progress made in several areas. Under the first two thrusts, considerable attention has been devoted to defining and understanding the operational decision environment, to understanding the nature of stressors that operate in this environment, and to conceptualizing and developing measures of tactical performance at the individual and team levels. In addition, a number of approaches for training individuals and teams to perform effectively in stressful tactical decision making tasks have been identified for further development (the third thrust listed above). Finally, under the fourth and fifth thrusts, initial principles of decision support and display have been specified for follow-up. In the sections that follow, research directions associated with these areas will be described, along with a brief delineation of future plans in the area.

UNDERSTANDING THE DECISION ENVIRONMENT

From a theoretical standpoint, relatively little attention has been paid in past work to the kind of tactical decision making that occurs in a dynamic, fast-tempo tactical environment (Orasanu and Salas, in press). Instead, there has been a tendency to study decision making under part-task and static conditions. The literature does, however, provide a solid foundation upon which to base a systematic investigation of tactical decision making. In fact, several recent developments in decision making theory (and related fields) provide strong leads which require elaboration and validation.

For example, recent advances in theories of human cognitive performance (Klein, 1989) provide useful extensions to traditional decision making theory. According to these recent approaches, expert decision makers may rely on well-developed memory representations to guide decision making in new (but similar) situations. From this point of view, the decision making process is considered to be intuitive rather than analytical. Other views of decision making hold that the process consists of a set of relatively discrete stages or sub-processes (e.g., Janis and Mann, 1977; Phelps, Pliske and Mutter, 1987). The primary advantage of these "stage models" is that, by separating decision making behavior into components, they provide a basis for studying the various mental activities associated with decision making and skills required to perform them successfully.

While the stage model approaches to decision making describe the phenomenon in somewhat generic terms so that lawful relationships or general characteristics of decision making can be revealed, the notion that decision making processes are situation-specific is also fairly well accepted (Ebbesen and Koneci, 1980; Sage, 1981). Many theorists postulate that characteristics of the decision situation and the nature of the decision itself influence how decisions are made (and, more importantly, how they should be made). Others have theorized that characteristics of the individual decision maker(s) have a significant impact on the decision making process; this condition poses the interesting challenge of investigating the merits of adaptive and prescriptive decision aiding features.

Past research has demonstrated that a host of situational and individual variables are important factors in decision making. The value of such research is that it highlights the complexity of naturalistic decision making and has the potential to provide guidance regarding factors that must be considered in studying decision making under particular operational conditions. Unfortunately, the state of the literature is such that it lacks integrating or organizing principles. Experiments have been conducted in numerous disciplines with diverse variable definitions and methodologies, making it difficult to draw conclusions across studies. Further effort is needed to organize this literature in a manner that will allow determination of: a) which factors are critical to decision performance in the operational environment, and b) the manner in which such factors affect individual and team decision making in this context.

From a descriptive standpoint, progress has been made in TADMUS in defining the operational decision environment via extensive interviewing and observing of actual Navy crews. The following characteristics appear to be present in the operational environment:

- Multiple information sources
- Incomplete, conflicting information
- Rapidly changing, evolving scenarios
- Requirement for team coordination
- Adverse physical conditions
- Performance pressure
- Time pressure
- High work/information load

- Auditory overload/interference
- Threat of hostile engagement

A second benefit of these interviews and observations was to determine which stressors are likely to have an impact on decision making performance. Of these, it was necessary to select a sub-set of stressors that could be induced reliably and safely in experiments. It was decided that the following stressors are candidates for manipulation in experimental studies:

- Task-related stressors:
  - Workload/time pressure
  - Uncertainty
  - Auditory overload

- Ambient stressors:
  - Auditory interference
  - Fatigue/sustained operations

Once potential stressors for manipulation were identified, a survey of the literature was conducted to determine the most common ways these stressors have been manipulated in past research. Manipulation techniques were assessed on the basis of: reliability with which stress could be created, the face validity of the procedure, ethical concerns, safety and feasibility. Based on this review and interaction with fleet personnel, the following initial stress manipulations were established:

1. *Workload/time pressure*. Workload and time pressure will be manipulated via the tactical task (problem) scenario. Specifically, the number of tasks an operator must simultaneously perform will be increased as a manipulation of workload. Time pressure will be defined as the time allowed to prosecute a contact and manipulated by changing the distance at which a potentially hostile contact is inserted into the scenario (i.e., the closer the contact is to own ship upon initial detection, the less time the team will have to prosecute it).

2. *Uncertainty*. Uncertainty will be defined as the amount of information available regarding a contact's intent. Under high uncertainty conditions less information will be available as, for example, when an identification mode is not available on a contact.

3. *Auditory overload*. Auditory overload refers to an overload of information being received via auditory channels. In actual CICs, operators are required to monitor several channels of auditory information simultaneously (both intra-ship and external communications) in order to receive information that is crucial to their task. Auditory overload will be manipulated by creating scenario-specific tapes that can be fed into operator headsets during an exercise. Tapes will include information that is pertinent to the task and scenario events.

4. *Auditory interference*. In contrast to auditory overload, auditory interference refers to excessive background noise that creates confusion or distraction, but is not necessarily germane to the operator's task. It will be manipulated in TADMUS experiments by taping actual CIC conversations that are unrelated to the identification problem (e.g., conversations between the air intercept controllers and pilots), and then playing these tapes as background noise during scenarios.

5. *Fatigue/sustained operations.* Results of fleet interviews and surveys indicate that sustained operations are a frequent stressor in CIC. Based on past work, sustained operations will be manipulated by extending the time operators must work. The details of this manipulation will be finalized later in the project.

Future Plans

A series of baseline studies will be conducted to document the impact of stressors on tactical decision making performance. These experiments will employ measures of performance (described below), and will manipulate several operationally relevant stressors as delineated above. Results of these experiments will provide data regarding the expected performance decrements under stress, clues as to how to ameliorate the negative effects of stress, and a baseline of performance under stress to which later results can be compared.

MEASURES OF TACTICAL PERFORMANCE

In order to assess the efficacy of interventions designed to improve tactical decision making performance, it must be possible to measure performance in a valid, reliable manner. Furthermore, measures that indicate *why* particular performance outcomes occurred can provide information that can be fed back to improve the intervention. While these statements may seem obvious, it is most often the case that operational systems lack a viable performance measurement system. This is particularly true in complex systems where it is difficult to identify a single correct answer, or where multiple operators are involved. It was necessary in TADMUS, therefore, to develop a performance measurement system for tactical decision making.

To begin with, an overall approach was adopted to guide performance measurement development. This approach assumes that there are different classes of measures appropriate for different purposes, and that no single measure of performance will be appropriate for all purposes. Specifically, *descriptive* measures seek to document behavior in individuals and teams, highlighting crucial points of interaction and moment-to-moment changes in team functioning. Such measures provide a foundation of understanding about what happened in a decision making scenario, and are a precursor to all other measures. Given the complexity of CIC decision making, even this level of measurement (i.e., simply describing what is happening at any given time) can be difficult to accomplish.

The second level of performance measures is *evaluative*. Such measures assume that there are identifiable standards of performance against which current performance can be judged. Evaluative measures answer questions regarding the effectiveness of a decision strategy or course of action (e.g., quality, timeliness, accuracy, etc.). These measures can be normative, or based on expert judgement.

Finally, *diagnostic* measures seek to identify the causes of behavior. They focus primarily on how and why things occurred as they did, and form the basis of feedback for subsequent performance improvement. As such, they are paramount to training and to system improvement.

It has also been useful to distinguish between *outcome* and *process* measures. Outcome measures describe what happened as a result of performance--decision accuracy, latency, quality, or timeliness; number of contacts correctly identified; time to detect a contact, etc. Process measures focus more on *how* the task was accomplished. They describe, for example, the interactions among team members, information flow during various phases of the task, and task strategies employed by team members.

They are crucial to understanding and determining the causes of performance outcomes, and to provide a basis for diagnosing performance.

Finally, effort under TADMUS has been directed at developing measures of team performance, as well as measures of individual performance. The sections that follow expand upon our rationale for measuring tactical performance, with emphasis on team-level measures being developed.

Team Performance Measures

Several researchers have suggested recently that teams comprise the cornerstone of modern industrial operations (Hackman and Morris, 1975; Cummings, 1981). It is not surprising, therefore, that work teams have been the subject of countless investigations over the past few decades (Cannon-Bowers et al., 1990). Despite this volume of research, however, relatively little is known about fundamental processes underlying teamwork and coordination, or how best to train teams to perform effectively (Salas et al., 1985). In particular, past research has produced useful descriptive models. More work is needed to develop explanatory models and measurement techniques in order to understand how teams acquire, maintain and lose critical teamwork skills.

Recently, a series of studies conducted with military command and control teams and aircrews has made significant progress in understanding the nature of teamwork (Glickman et al., 1987; Oser et al., 1989; Stout et al., 1990). To begin with, Glickman et al. (1987) found that two separate tracks of behavior evolve during team training. The "taskwork" track involves skills that are related to the execution of the task and/or mission (e.g., operating equipment, following procedures). The second track, labelled the "teamwork" track, involves skills that are related to functioning effectively as a team member. To summarize the overall findings of this and related work, the following conclusions can be drawn:

a) Behaviors that are related specifically to team functioning (i.e., independent of the particular task at hand) are important to task outcomes (Oser et al., 1989; Stout et al., 1990).

b) Effective teamwork behavior among skilled teams appears to be fairly consistent across tasks (Oser et al., 1989).

c) Team process variables (e.g., communication, coordination, mutual performance monitoring, compensatory behavior) influence team effectiveness (Stout et al., 1990).

In terms of specific teamwork behaviors, McIntyre et al. (1988) recently summarized the lessons learned from investigations of tactical teams and concluded that teamwork appears to be comprised of a complex of behaviors including: closed-loop communication, compensatory behavior, mutual performance monitoring, giving/receiving feedback, adaptability and coordination. Further, McIntyre suggests that in effective teams, members seem to be able to predict the behavior and needs of other members.

There is an important implication derived from research described above for diagnosing and measuring team performance. Specifically, evidence presented above suggests that process measures are a necessary adjunct to task outcome measures (i.e., those that are related directly to the goals of the task, such as number of targets successfully prosecuted, and deviations from specified course) for assessing team performance. This is particularly true when the purpose of measurement is to provide feedback that will improve performance (Coovert, Cannon-Bowers and Salas, 1990). Outcome measures do not contain information that is useful for diagnosing the cause of poor performance, or describing how team members should adjust behaviors to affect improved performance in the future. For example, informing a team that they have misclassified a contact does not provide guidance that will enable them to perform successfully in the future. On the other hand, informing them that their error was due to a specific breakdown in communication among two members (i.e., diagnosing and communicating the cause of the poor performance) gives them insight into how subsequent performance can be improved.

To date, several studies have indicated that communication behaviors and other process measures are related to team effectiveness (Stout et al., 1990; Lassiter et al., 1990; Oser et al., 1989). In addition, research has produced team performance observational scales (Morgan et al., 1986; Glickman et al., 1987). Using a critical incidents technique, these researchers interviewed team training instructors in order to generate behavioral examples of effective and ineffective teamwork. These items were then included as part of a checklist that was used by training instructors to rate teams in training. Evidence from these studies indicated that effective and ineffective teams could be distinguished based on the frequency of behaviors exhibited in various categories on the scale.

More recently, observational measures of team performance have been developed to assess the teamwork skills of cockpit crews (Franz et al., 1990). Based on a needs assessment conducted with helicopter pilots, an instrument is being developed that requires an observer to rate crews on critical teamwork skills that have been grouped into several behavioral dimensions. This is slightly different from the Glickman et al. (1987) approach of recording behavioral frequencies. Preliminary evidence suggests that such a methodology can be successful in distinguishing effective from ineffective teams (Stout et al., 1990).

Other researchers have been concerned with developing measures of team communication as indicators of effectiveness. With respect to communication patterns (i.e., the form or structure of a series of interactions), several investigators have found that overall frequency of communication among aircrew members is directly related to crew effectiveness (Foushee and Manos, 1981; Foushee et al., 1986). In addition, communication content has also been found to be related to aircrew effectiveness. For example, it has been found consistently that the frequency of commands, observations, suggestions and acknowledgements is positively related to performance (Krumm and Farina, 1962; Foushee and Manos, 1981). One of the goals of the TADMUS program is to use these findings to develop training strategies to improve communications techniques (e.g., more effective use of brevity codes) among CIC team members.

Progress has been made under TADMUS in the development of a scale to record and assess teamwork and team process. Based on work described above, a modified critical incidents technique was used to develop initial scales that will enable both experimenters and instructors to record the quality and quantity of teamwork behaviors exhibited by the team. It will be useful as a means to evaluate teamwork skills in tactical environment, and will enable researchers and instructors to infer the causes of poor or ineffective performance in teams. As such, it falls into both the evaluative and diagnostic categories delineated above.

Development of the scale was accomplished by having subject matter experts at operational sites generate initial critical incident examples for all positions in the anti-air warfare team. These incidents were cast into a format to allow assessment of critical teamwork skills (including mutual performance monitoring, compensatory behavior, giving/receiving feedback, team orientation, communication and coordination of activity). Initial validation of the scale will be accomplished by testing it using high fidelity simulations as well as in laboratory experiments. Scale format will also be adjusted according to input/reaction from fleet personnel. Two versions of the scale will be available when validation work is complete: one for use by experimenters in baseline and intervention experiments and a second version for use by instructors in training situations.

Human Performance Modeling

While the observational scales and communication assessment techniques described above are useful as indicators of team performance, other measures of team effectiveness must be developed to augment these. Given the rapidly changing, complex, dynamic tasks faced by many military teams, a more fine-grained analysis technique is needed that can reflect moment-to-moment variations in performance. This is critical particularly to the development of feedback mechanisms, which must allow for delivery of timely, accurate performance information as tasks are being completed. Therefore, it is necessary to develop methods to measure behavioral latencies, behavioral probabilities, distribution of workload across members, task sequencing, and other aspects of team process. One way to accomplish this is to employ human performance modeling techniques (see McMillan et al., 1989). In general, such techniques have the advantage of being able to describe complex, multi-operator tasks in a manner that is amenable to analysis. Further, they can provide a basis for measures of performance by specifying in detail what is to be considered effective team performance. A methodology that holds particular promise in this regard involves the use of Petri nets; this technique will be described in the following sections.

Petri Net Modeling. Petri nets are the basis of a modeling technology that has been used to model a variety of processes (Levis, 1989; Reisig, 1985). In general, Petri nets specify the relationship between two types of entities, *places* (or states) and *transitions* (events). Graphically, places and transitions are represented as circles and rectangles, respectively, and are connected by directed *arcs*. Information, data or conditions are represented as *tokens* in a Petri net. Tokens reside in places and move from one place to another through the firing of the transitions which connect the places.

As a modeling tool for teams, Petri nets have several attractive properties (Coovert et al., 1990). The first is their ability to model two critical aspects of team performance, *concurrency* and *conflict*. Concurrency occurs when different team members perform individual tasks independently (at the same time), which is the case with most military tactical teams. Conflict occurs when more than one team member requires a shared resource, or when the activity of one team member interferes with the activities of another.

The second desirable property is that Petri nets are useful for modeling at various levels of abstraction. It is possible to construct a net to represent each individual (i.e., an individual's information processing actions), so that the individual would be the unit of analysis. It is also possible to construct a team-level net, where each place represents a team member, or a high-level net representing an entire organization. Places in such a net could represent, for example, functional departments. Furthermore, the arcs in a Petri net could easily represent serial and/or parallel interactions among members, subteams, or teams. Once linkages are established, one can examine the influence of constraints at various levels of abstractions (e.g., individual, subteam, team, department or organization).

The third desirable property of Petri nets is that models may be analyzed in a variety of ways to validate a theoretical model of team performance and gain insight into the behavior of the system. They allow a rigorous examination of how teams behave in pursuit of task accomplishment. They can be verified via comparison of different team nets to one another, and validated via comparison of various net configurations with external indices of performance (e.g., outcome measures, expert ratings).

Petri nets can be employed to study team performance by constructing a functional net of the team task. A "functional task net" describes in a step-by-step fashion how the task is executed, including the interaction among team members. Once the functional net is constructed, other nets representing such things as communication patterns, individual decision making, and resource transfer can be "layered" on top of the functional net. That is, using the functional net as a foundation, other nets can be developed that describe only selected aspects of team performance. These nets are linked to the functional net because they are all based on a common task structure, but they are simpler than the functional net since they represent only a subset of tasks. For example, a verbal communication net could be devised that represents only the verbal communication among team members (i.e., it would contain only those tasks from the functional net that require verbal communication among team members). These specialized nets allow a focused analysis of various aspects of team performance. In addition, the functional net itself can be studied to observe how it changes as a function of such variables as team member experience, workload, time pressure and other types of stress.

In terms of TADMUS progress, a laboratory study was conducted to determine the appropriateness of Petri nets for modeling team tasks. Results of this effort indicated that functional models of team performance were sufficiently detailed to describe team performance. Further, it was found that several net-based measures (e.g., times a transition fired, time spent at particular transitions) could distinguish between effective and ineffective teams (Coovert et al., 1990).

Data from the lab study provided a basis upon which to proceed with net development in operational settings. Operational training and systems manuals were used as a basis to specify initial team nets for an anti-air warfare task. In parallel, nets were built in conjunction with personnel at Navy training facilities. Initial nets were then upgraded and refined based on input from fleet personnel. They will be tested using experienced teams engaged in high-fidelity exercises at training centers and on ships. Specifically, net-based recordings of team performance will be used as a basis for the decomposition of complex tasks, for understanding inter-relationships between tasks and for the development of improved training strategies.

Future Plans

Team and individual performance measures developed under TADMUS must be tested and refined. Initial empirical work will be devoted to this task, so that measures may be employed in later experiments and incorporated into training system design.

TRAINING AND SIMULATION PRINCIPLES

A third major thrust of research underlying the TADMUS project deals with developing training and simulation principles that help to make decision makers less susceptible to the adverse effects of stress. Our approach to this task has been to specify three broad categories of interventions that are hypothesized to be effective:

Increase overall readiness. It can be hypothesized that individuals and teams that are highly competent in executing a task will be more resilient to the negative effects of stress on performance. Therefore, a goal of TADMUS is to improve methods to train tactical decision making in teams and individuals, so that the likelihood of reaching task proficiency is enhanced, and the amount of time required to reach proficiency is reduced.

Expose trainees to stress during training. Several theorists have suggested that stress exposure training can inoculate them from the impact of stress in the actual task performance situation (e.g., Novaco, 1983; Meichenbaum, 1985). Methods to implement stress exposure techniques in a tactical decision making environment will be investigated in TADMUS

Target skills that are particularly vulnerable. Based on results from baseline studies (described above), it may be possible to identify skills that are particularly vulnerable to the effects of stress. An effort will be made to concentrate attention on these skills in training, so as to reduce the impact of stress on tactical decision making performance.

To date a number of potential training strategies have been identified that appear to warrant follow-up. The following paragraphs will describe in greater detail the theoretical foundation, rationale, and hypotheses for several of these potential training technologies.

Shared Mental Model Training

The notion of mental models has been invoked as an explanatory mechanism by those studying skilled performance and system control for a number of years (Veldhuyzen and Stassen, 1977; Jagacinski and Miller, 1978; Rouse and Morris, 1986). According to Rouse and Morris (1986), a mental model can be defined as a "mechanism whereby humans generate descriptions of system purpose and form, explanations of system functioning and observed system states, and predictions of future system states".

In the area of cognitive psychology, researchers have suggested that mental models are important more generally to the understanding of how humans interact and cope with the world (Rouse and Morris, 1986). For example, Williams et al. (1983) maintain that mental models allow people to predict and explain system behavior, and help them to understand the relationship between system components and events. Wickens (1984) contends further that mental models provide a source of people's expectations. In an even more general view, Johnson-Laird (1983) suggests that people "understand the world by constructing working models of it in their mind". Mental models enable people to draw inferences and make predictions, to understand phenomena, to decide what actions to take, to control system execution, and to experience events vicariously (Johnson-Laird, 1983).

In reviewing the literature pertaining to mental models, Rouse and Morris (1986) concluded that a number of common themes can be drawn among theories that describe the purpose of mental models--namely that mental models serve to help people describe, explain and predict system behavior. It must also be noted that most theorists conceptualize mental models as more than simple mental images. Instead, mental models are manipulable, enabling people to predict system states by mental manipulation of model parameters (see Johnson-Laird, 1983 for a detailed description of mental model functioning). Klein (1989) has suggested, for example, that expert decision makers engage in a mental simulation that allows them to predict the ramifications of a potential decision prior to taking action.

With respect to training, a number of theorists have hypothesized that training that fosters development of accurate mental models of a system will improve performance. According to Rouse and Morris, for example, one of the purposes of instruction is to develop mental models necessary to execute the task. Recent evidence suggests that the manner in which people cognitively structure information about a task has an impact on the way new information is assimilated and learned (cf. Eberts and Brock, 1987). Information that is compatible with existing mental models will be easier to learn (Wickens, 1984). Pre-existing models of the task can also have an impact on training. Rouse and Morris (1986) contend in this regard that incorrect mental models can impede learning. Furthermore, evidence suggests that pre-existing models may be difficult to eliminate; they appear to persist in the face of correct information (DiSessa, 1982). Finally, evidence suggests that to be most effective and generalizable, training must provide a conceptual model of the system, along with specific guidance or cueing in how general principles of system functioning are applied (Rouse and Morris, 1986).

Team Performance and Mental Models. Research cited above provides support for the contention that teamwork behaviors can be isolated from other task-related behaviors. In terms of training requirements and strategies, further research is needed to translate identified teamwork behavioral dimensions into requisite knowledge, skills and abilities (KSAs). For several classes of teamwork behavior such as communication, giving and receiving feedback and mutual performance monitoring, KSA development seems to be fairly straightforward.

It is in the area of defining and training skills associated with *coordination of action* and *adaptability* that little is known. This is due to the fact that these skills appear to involve the ability of team members to predict the needs of the task and anticipate the actions of other team members in order to adjust their behavior accordingly. In other words, team members appear to form expectations of what is likely to happen. Particularly when a novel situation arises, teams that cannot strategize overtly must exercise shared or common models of the task and team in order to maintain performance (Kleinman and Serfaty, 1989; McIntyre, et al., 1988; Orasanu, 1990; Cannon-Bowers, Salas and Converse, 1990). The role of mental models in explaining team behavior, then, stems from their ability to allow team members to generate predictions about task and team demands in absence of communication among team members.

The notion presented thus far of team mental models and how these relate to team effectiveness has several implications for the understanding of team performance and training. As an explanatory mechanism, the team mental model construct is useful in understanding how teams are able to coordinate behavior and select task strategies in absence of explicit coordination activities. Under conditions of high workload, time pressure and other kinds of stress, such implicit coordination appears to be critical (Kleinman and Serfaty, 1989).

With respect to training, the shared mental model idea suggests that training strategies designed to foster development of shared mental models has the potential to improve team performance. Research cited earlier regarding the success of efforts to train mental models for system operation offers preliminary evidence that such training may be possible. For example, research suggesting that particular knowledge structures (i.e, mental models) can be trained provides support for the notion that common expectations for the task and team can be developed through training (Cannon-Bowers and Salas, 1990)

From what has presented to this point, it may be hypothesized that specific training strategies which may be useful in training shared mental models include:

1. *Positional clarification.* Interventions designed to provide information regarding the structure of the team and task, the interrelationships among team member positions, and the roles and responsibilities of each team member could be hypothesized to improve team performance by enhancing common task and team expectations. Such training, which represents requisite team and task knowledge, could be presented via lecture, computer assisted instruction, or via written material. Such training would represent initial preparatory training, but would probably not be sufficient to develop shared mental models. Another potential training technique that may be useful for this purpose is role playing, which also has the benefit of making the trainees more active participants in the training.

2. *Guided practice and feedback.* Past research has indicated that unguided practice can lead to development of inaccurate mental models. These findings suggest that teams should practice tasks under the guidance of instructors. Feedback and debrief mechanisms must be designed to result in accurate, common expectations for the task and team. In addition, feedback regarding *specific behaviors* that must be changed should be more effective in establishing accurate expectations than general, less specified feedback. Simulation facilities would be a most appropriate means to provide such practice opportunities. Recent evidence with aircrews suggests that low-fidelity simulation may also be viable (Stout et al., 1990).

3. *Cross training.* A potentially useful strategy to train common mental models may be to cross train team members on tasks that are related to their own task. Such training would be beneficial to the extent that it helps team members to learn what their teammates will need (in terms of resources, information, assistance), given various task demands.

4. *Instructor training.* Much of the success of team tactical training depends on the quality of instructors. With respect to shared mental models, instructors must be trained to recognize effective teamwork behaviors and other evidence that team members share common mental models as a basis to deliver feedback.

5. *Team leader training.* Training team leaders to foster development of shared mental models also has potential value. It can be hypothesized that team leaders who are trained to articulate their own view of the task and team, who encourage discussion and strategy formation among team members, and who make clear their expectations of team member behavior should be successful in helping their teams to develop shared mental models

Other Training Strategies

Several other avenues are being pursued under TADMUS with respect to developing training and simulation principles. For the most part, these are in the early stages of development, and will be described only briefly here.

The theory of controlled versus automatic processing (see Shriffin and Schneider, 1977; Fisk, Ackerman and Schneider, 1987) provides a basis upon which to hypothesize that overtraining may be a viable means to help individuals be more resilient to stress effects. The theory postulates that certain task components can be "automatized"--that is, performed with little demand on cognitive resources. Automatized skills can be performed rapidly and effortlessly even under extremely exceptional conditions such as stress. The implication of this research for tactical decision making training is that it may be possible to train certain components of the decision making task to automaticity so as to free up critical resources for higher order cognitive functioning. This is particularly critical in stressful or emergent situations. In terms of training system design, this research can lead to principles for overtraining decision makers, and for part-task training (i.e., training only selected aspects of the larger task).

As mentioned above, several researchers have suggested that exposing people to stress in training may inoculate them from the effects of stress in task performance (Novaco et al., 1983; Zakay and Wooler, 1984). While much of the research in this area has been conducted in clinical settings, there are clear implications for TADMUS. Briefly, several underlying principles of stress exposure training require further attention. These include: the role of self-efficacy and confidence in resilience to stress; the impact of shifting attentional focus away from the stressor and on to the task; and the impact of stress exposure on transfer of training. Research into these factors will help to specify whether stress exposure training is a viable strategy for tactical decision making tasks, and if so, how it should be designed and implemented.

Finally, the theory of pattern primed or schema-based decision making (described in detail in the next section) may provide an important basis upon which to design training. Briefly, the theory holds that decision makers rely on previously established memory structures in order to make decisions in the current situation (Klein, 1989). In terms of training system design, the theory raises a number of important questions regarding how to structure training exercises and scenarios so as to optimize training value. These questions include: (1) What are the implications of adopting a schema-based theory of decision making for training system design? (2) What types of scenarios are needed to train in a schema-based strategy? (3) What are the critical cues decision makers use in recognizing schema and making decisions? (4) How are various cues mapped to decision outcomes (both in terms of the process by which cues become associated with outcomes and in terms of how particular cue-outcome patterns can be identified)? (5) How should training exercises be designed so that they incorporate critical cues, and impart knowledge regarding the relationship between particular cues (or cue patterns) and decision behavior? (6) What kind of feedback needs to be given in training to maximize training success? (7) How wide a range of scenarios do decision makers need to be exposed to? (8) What factors in training facilitate the ability of decision makers to recognize and respond to critical cue patterns so as to improve decision making performance? The answers to these questions will provide insight into how to optimally structure and sequence training scenarios.

Future Plans

Once initial training and simulation principles have been established based on theories described above, specific hypotheses will be generated and tested in laboratory and field settings. The outcome will be a series of training and simulation principles for application in stressful environments.

DECISION SUPPORT PRINCIPLES

Research in decision making has shown that even under virtually perfect conditions, humans (including experts) have biases in judgment and may make consistent errors (Tolcott, 1991) [published in this issue of ESNIB]. Under suboptimal conditions, the potential for human error is even greater. Events have shown that decision errors can have serious political, as well as tactical, ramifications and are often scrutinized in great detail long afterwards. Consequently, the need to improve decision making via enhanced decision support and displays is more critical than ever. Unfortunately, a recent workshop (Tolcott and Holt, 1987) found that "although research on human judgment and decision making has produced much information and many theories about cognitive abilities and fallibilities, these results have rarely been considered in decision aid development. Other conferences on decision support have come to similar conclusions. TADMUS research in this thrust is intended to apply lessons learned in decision research to the development of decision support and information display principles for stressful environments.

Decision Making and the Design of Decision Support Systems

In the past five years, there have been very significant changes in the direction of research and thinking about how people make decisions. Early theories of decision making often tried to create a normative, or rational, model of how people reason, perhaps as a way to describe how they ought to reason (Tolcott, 1991). More recently, however, it has become clear from the work of several researchers, notably Klein (1989) and Noble (1989), that, in naturalistic situations (characterized by uncertainty, ambiguity, high stakes, action/feedback loops, stress, etc.) people actually use fairly straightforward, perhaps simple decision making strategies that make effective use of their substantive knowledge without overtaxing their capacity (Cohen, in press). Moreover, Payne and Bettman (1988) have demonstrated that decision makers will actually change strategies based on the characteristics of the problem they face, providing evidence that a model which relies on a single strategy is insufficient to explain the complex cognitive processes that people use (even if the strategy is simple).

One of the most important enhancements to understanding began with Simon's theory of satisficing (Simon, 1957), which proposed that decision makers, limited by their ability to analyze a problem exhaustively, use shortcuts to reach decisions and judgments efficiently and with satisfactory accuracy. Klein's research carries this even further by suggesting that, for naturalistic decision problems, decision makers select the first response that they imagine will satisfy the problem at hand. The implications of these ideas is that efforts to support tactical decision making by providing the means for the decision maker to generate and compare all possible options for the purpose of optimizing his decision may well be misguided. The TADMUS program seeks to explore the application of these new ideas about how people actually make tactical decisions to decision support system design.

Apart from these findings that people do not use normative strategies to make tactical decisions, there are practical problems that limit the ability to develop classical option generation/comparison decision support systems. In fact, many of the factors that make it difficult for people to use a systematic reasoning process make it difficult for computers as well. For example, the data required for making optimal decisions is often missing, ambiguous, unreliable or late; there may not be enough time to make the necessary calculations; the consequences of an action may not be known or knowable, making risks uncertain; and it may be necessary to employ an incremental process where action/feedback loops are required to gain additional knowledge of the situations. Consequently, the TADMUS program has taken the position that the best approach to improving decision making in the naturalistic environment is to better understand and support the strategies that people actually use to solve problems, rather than attempt to supplant them.

In an effort to understand better how people actually make decisions in the complex environment of Navy ship self-defense, the project has adopted a multi-faceted approach. In one effort, the literature was surveyed in order to establish a comprehensive set of distinct decision strategies which researchers have postulated and described these in formalized terms. Concurrently, many structured interviews have been conducted, using a critical incident analysis technique, with Naval officers who have been engaged in the kinds of decisions of interest. At present, interview data are being evaluated and analyzed systematically to determine which, if any, of the formalized decision strategies were utilized by the Naval officers, to identify other strategies which appear to have been used and to characterize the conditions which elicit the different strategies.

Although this effort is incomplete at this time and preliminary judgments are subject to change, several findings are beginning to emerge which support the multi-faceted model or provide new insights. For example, it appears useful to divide the tactical decision process into two cognitive activities: situation assessment and response selection. Early data suggest that decision makers struggle to make sense of an unfolding situation either by attempting to match the series of events they are witnessing with similar generalized memories of experiences they have had before, or else by creating plausible stories about how the events they are witnessing could be happening (e.g., what the opponent's intent is). Once an assessment of the situation has been formed (and perhaps believed) the most salient response is often readily matched to a memory of a similar experience and is applied through a process of mental simulation (Klein, 1990). If the first response choice does not pass the simulation test, another response choice may be mentally simulated, and so on.

A key idea to the process described here, as compared to the rational model, is that responses are considered sequentially rather than in parallel. Past work in support of the two-stage model (situation assessment and response selection) of decision making by Klein (1989) has found that decision makers often claim that they are not aware of making decisions. They simply know which response to make in highly stressful environments based on their training and experience. He has termed this decision strategy recognition primed decision-making (RPD) because it rests on the ability to recognize a familiar pattern in what, to novices, may appear to be a complex, ambiguous situation.

The RPD model describes a highly iterative process of decision making. Incremental actions/responses can be taken to elicit responses from a threat in order to understand the situation better. Certain events, occurring within certain time periods, would be expected based upon a mental simulation. These events, when observed, serve to improve pattern matches or reassurance that the storyline is correct. This model clearly has many implications that are important for decision support, training and the man-machine interface -- the thrusts of the TADMUS program. With respect to the RPD program and its implementation, three of the most important questions are: (1) How can the situation assessment strategies for matching events to experience and explanation of events be duplicated or supported by computer? (2) What is the role of expertise and training? (3) How can the computer be used to support mental simulation for response selection?

In an effort to apply the implications of the RPD model to decision support, the TADMUS program is exploring the work of Noble, who has been following a line of research similar to RPD, based on the idea of cognitive schema. Noble (1989) has spent the last several years in developing computer-based situation assessment support for tactical decision making and is developing this capability further to support real-time decision making using techniques for situation templating. In the TADMUS Decision Support System, situation templating will comprise one of several tools available for situation assessment. Using the situation templating tool, a decision maker will be able to plan and train for expected situations that might occur at any moment using his own storehouse of knowledge and experience. Situation templates specify the kinds of participants and activities expected in a situation. Each template also specifies the observable manifestations of these participants and activities. According to Noble's version of RPD:

1. Salient situation features "activate" in memory the most relevant of the general situation templates to create initial hypotheses about possible ongoing and future activities. These hypotheses include specific expectations about the situation, including the types, times, places, performers and indicators of expected events.
2. Indicators able to confirm or contradict situation expectations are identified and sought.
3. Successfully-retrieved indicators are used to test the expectations of each of the situation hypotheses.
4. Hypotheses not consistent with retrieved indicators are eliminated. Hypotheses that are consistent are updated, refining the hypotheses' expectations.

The result of this situation assessment process is one or more candidate situation hypotheses, each associated with a degree of belief and a set of the expectations about the types, times, locations, and performers of events. The degree of belief reflects the extent to which observed situation data confirm or contradict the situation expectations.

Situation uncertainty is an inherent part of the assessment, and arises whenever the situation data cannot identify the precise characteristics of the situation. Noble distinguishes between two different types of uncertainty. The first type is uncertainty about the *kind of situation*. Such uncertainty arises when no single situation model accounts for the available information much better than any other situation model. Noble represents this type of situation uncertainty as multiple situation hypotheses. The second type is uncertainty about the *precise characteristics of the expected events or participants in a situation*. This type of uncertainty arises when the available situation data and the situation models in memory cannot define fully the situation characteristics. Noble represents this type of uncertainty as bounds on the values of these characteristics.

Besides a recognition based strategy, a second strategy for assessing situations that has so far been identified from the research literature and the structured interviews has been generally referred to as explanation-based reasoning (Pennington and Hastie, in press). How, or even whether, this strategy differs fundamentally or qualitatively from RPD or situation templating is still a topic of debate among the program participants. At this point in the program, it has been decided to treat it differently from an implementation point of view, whether or not it relates to a different cognitive process. Consequently, a second decision support tool to become part of the TADMUS decision support system is under development as Situation Assessment By Explanation-based Reasoning (SABER).

Explanation-based reasoning is used in SABER as a means of deciding which possible conclusion is best supported by a set of data that may be incomplete or contradictory. An explanation in this framework is simply a causal model that incorporates all available data into a coherent structure that supports one of several situation assessments. In SABER, explanations are evaluated according to three primary criteria: simplicity, completeness, and data weights. Simplicity relates to how well an explanation accounts for a conclusion without using additional assumptions. Completeness considers how much of the available data is directly accounted for by the explanation. Each type of data is weighted so that the relative importance of different kinds of data can be considered.

Two other kinds of weights are also considered. Where contradictory data are involved, weights assigned to the different assumptions used to explain contradictions are used to test the plausibility of applying those assumptions. Where certain kinds of data are expected to occur in support of a given conclusion, but have not been observed, negative weights are applied to decrease the degree of belief in that conclusion. Before SABER can construct explanations, the set of possible conclusions and the set of possible kinds of input data must be specified. The representation of each data type is itself a set of explanations. The primary explanation associated with a data item indicates how the given kind of data supports one particular conclusion by default. Alternative explanations are attached to each data item to indicate the assumptions that could be used to override the default conclusion.

When several kinds of data are present, the individual explanations for each particular kind of data are combined into larger explanatory structures. Each such structure can be shown to the decision maker, including: (l) the complete set of input data, (2) the conclusion each explanation points to, (3) an indication of the strength of the conclusion, and (4) a list of evidence for the conclusion. The strength of a conclusion can be indicated by confidence levels: confirmed, probable, possible. Where data are contradictory, certain explanations can be indicated as preferred. However, which information to display and how to display it are issues yet to be addressed in this effort (see the discussion regarding information display below). One useful characteristic of this tool is that it accounts for the differences in importance of various data. For instance, a visual identification of an aircraft will probably carry more weight than electronic indications that suggest another identification.

Other tools, in addition to the RPD and SABER, may be suggested as evaluation of TADMUS data continues. Translating strategies that may yet be identified as important to support into implementable computer-based decision support tools is by no means straightforward; it requires considerable insight and creativity. Furthermore, several critical research issues have only begun to receive attention and will be the subject of future reports. These questions concern how a decision maker decides how to decide (i.e. which strategy to use), how errors can be recognized, and the role of cognitive heuristics/biases in decision making and display design (Kahneman, Slovic and Tversky, 1982). Moreover, integrating a set of tools, driving them with real-time data, and providing a man-machine interface that permits the use of the tools under the effects of stress are some of the other developments being undertaken by the TADMUS program. Current plans for the implementation of the decision support system call for an architecture in which the tools, together with a display manager, communicate through a shared blackboard which is also served by existing tactical databases and communications links. Portions of this architecture have already been demonstrated and an early trial of the database connectivity to the situation templating system will be conducted shortly.

Future Plans

As the prototype decision support system is developed, it will be tested to determine its ability to improve decision making under a variety of operational conditions. Performance measures and stressors (described above) will be employed in studies to assess decision support system efficacy. Once the system is validated, principles will be extracted for application to other decision making domains.

DISPLAY DESIGN

The overwhelming majority of research on man-machine interface (MMI) design issues has been devoted to characteristics of displays that impact human perception, such as symbol legibility or detectability, and on relatively simple cognitive functions such as memory tasks. Few efforts have been devoted to understanding the effects of the style in which information is presented on more complex levels of human cognition such as decision making. Consequently, principles that can be applied to the design of the interface between the user and a decision support system for the purpose of enhancing cognitive processes, particularly in stressful, dynamically changing situations, are simply not available to any significant degree.

There is an urgent need for improved principles and models to guide the design of procedures for human-computer interaction and powerful computer interfaces for advanced combat and decision making systems. While shortcomings in the MMI design may not be fully evident during peacetime or non-critical events, it should be clear by now that these shortcomings can lead to catastrophic failures in a crisis. A technology base, which defines the effects of information presentation variables on decision making and which leads to guidelines for the design of displays for decision support systems is essential.

Several lines of investigation intersect in this task: relating higher level cognitive performance (i.e., decision making) to the way in which information is presented; relating the effects of stress to perception of displayed information and demonstrating that these effects can be managed; and relating team performance and training effectiveness to information display. What makes these efforts especially difficult to accomplish is that there is relatively little basic behavioral science in this area, with no generally-accepted theory of information transfer between a display and a user. Without such a theory there are few useful (standardized, accepted, sensitive) measures of performance with which to evaluate display design objectively.

In addition to what has been presented earlier in this paper regarding performance measurement, an approach to developing measures of performance that can be applied to display design may be to use the concept of information processing biases. It seems clear from behavioral experiments that people make consistent decision errors, but it is not as clear how to ameliorate them through display design (Tolcott, 1991) [published in this issue of ESNIB]. Effort in TADMUS is being devoted to determining whether particular display formats will decrease the likelihood of biased decision making.

Other evidence from cognitive psychology demonstrate that the manner in which a person structures information about a task impacts the way new information is assimilated and learned (see the previous discussion of mental models). Essentially, new information interacts with existing "mental models" of the world. Information displayed in a way that is compatible with such mental models is learned more easily and completely (Wickens, 1984). These research efforts will provide a basis upon which to explore the development of displays that are compatible with, and even enhance, the decision maker's cognitive representation of the problem and may suggest ways in which a commander can use information displays to enhance the team mental model.

Future Plans

Work in this area will include a series of analytical and experimental studies to examine decision processes, information processing biases, heuristics, and presentation formats. It will: (a) explore the effects of specific biases (e.g., representativeness, availability, anchoring, framing) under conditions of varying uncertainty and time stress; (b) investigate schema-based models as a basis for improved techniques for situation assessment; and (c) evaluate alternative display formats and techniques to counter the adverse effects of biases and stress.

CONCLUSIONS

Obviously, the challenge of improving performance in the highly-charged, complicated environment facing the modern tactical decision maker is a formidable one. The TADMUS program was designed to begin to answer critical questions regarding how human decision makers operate under exceedingly demanding circumstances, and how to improve such performance through training, decision support and display design. Principles developed and validated under TADMUS should have application to other environments that present similar demands on the human operators.

REFERENCES

Cannon-Bowers, J.A., and Salas, E. (1990). *Cognitive psychology and team training: Shared mental models in complex systems.* Symposium address presented at the annual meeting of the Society for Industrial and Organizational Psychology, Miami, FL.

Cannon-Bowers, J.A., Salas, E. and Converse, S.A. (1990). Cognitive psychology and team training: Training shared mental models of complex systems. *Human Factors Society, 32* (12), 1-4.

Cohen, M. (in press). In Klein, G., Orasanu, J. and Calderwood, R. (Eds.), *Decision making in action: Models and methods.* Norwood, NJ: Ablex.

Coovert, M.D., Cannon-Bowers, J.A., and Salas, E. (1990). *Applying mathematical modeling technology to the study of team training and performance.* Paper presented at the 12th Annual Interservice / Industry Training Systems Conference, Orlando, FL.

Cummings, T.G. (1981). Designing effective work groups. In Nystrom, P. C. and Starbuck, W. (Eds.) *Handbook of organizational design: Vol. 2.* London: Oxford University Press.

DiSessa, A.A. (1982). Unlearning Aristotelian physics: A study of knowledge-based learning. *Cognitive Science, 6,* 37-75.

Ebbesen, E.B., and Koneci, V.J. (1980). On the external validity of decision-making research: What do we know about decisions in the real world? In Wallsten, T. S. (Ed.) *Cognitive processes in choice and decision behavior.* Hillsdale, NJ: Erlbaum.

Eberts, R.E., and Brock, J.F. (1987). Computer-assisted and computer managed instruction. In Salvendy, G. (Ed.), *Handbook of human factors.* New York: John Wiley and Sons.

Fisk, A.D., Ackerman, P.L., and Schneider, W. (1987). Automatic and controlled processing theory and its application to human factors problems. In Hancock, P. A. (Ed.), *Human factors psychology.* New York: North-Holland Publishing Co.

Foushee, H.C., Lauber, J.K., Baetge, M.M., and Acomb, D.B. (1986). *Crew factors in flight operations III. The operational significance of exposure to short-haul air transport operations*. (NASA Technical Memorandum No. 88322). Moffett Field, CA: National Aeronautics and Space Administration.

Foushee, H.C., and Manos, K.L. (1981). Information transfer within the cockpit: Problems in intracockpit communications. In Billings, C. E. and Chaney, E. S. (Eds.). *Information transfer problems in the aviation system*. (NASA Technical Paper No. 1875, pp. 63-71). Moffett Field, CA: National Aeronautics and Space Administration.

Franz, T.M., Prince, C., Cannon-Bowers, J.A., and Salas, E. (1990, April). The identification of aircrew coordination skills. *Proceedings of the 12th Symposium Psychology in the Department of Defense*. (pp. 97-101). Springfield, VA: National Technical Information Services.

Glickman, A.S., Zimmer, S., Montero, R.C., Guerette, P.J., Campbell, W.J., Morgan, B.B., Jr., and Salas, E. (1987). *The evolution of teamwork skills: An empirical assessment with implications for training*. (Report No. TR87-016). Orlando, FL: Naval Training Systems Center.

Hackman, J.R., and Morris, C.G. (1975). Group tasks, group interaction processes, and group performance effectiveness: A review and proposed integration. In Berkowitz, L. (Ed.), *Advances in experimental social psychology: Vol. 8*. (pp. 45-99). San Diego: Academic Press.

Jagacinski, R.J., and Miller, R.A. (1978). Describing the human operator's internal model of a dynamic system. *Human Factors, 20*, 425-433.

Janis, I., and Mann, L. (1977). *Decision making*. New York: The Free Press.

Johnson-Laird, P. (1983). *Mental models*. Cambridge, MA: Harvard University Press.

Kahneman, D., Slovic, P., and Tversky, A. (Eds.). (1982). *Judgment under uncertainty: Heuristics and biases*. Cambridge, MA: Cambridge University Press.

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

Klein, G.A. (1990). Mental simulation and decision making. *Proceedings of the Fifth Mid-Central Ergonomics/Human Factors Conference*. Dayton, OH.

Kleinman, D.L., and Serfaty, D. (1989). Team performance assessment in distributed decision making. In Gilson, R., Kincaid, J. P. and Goldiez, B.(Eds.). *Proceedings: Interactive Networked Simulation for Training Conference*. Orlando, FL: Naval Training Systems Center.

Krumm, R.L., and Farina, A.J. (1962). *Effectiveness of integrated simulator training in promoting*

B-52 crew coordination. (MRL Technical Documentary Report No. 62-1). Wright-Patterson Air Force Base, OH: Aerospace Medical Research Laboratories.

Lassiter, D.L., Vaughn, J.S., Smaltz, V.E., Morgan, B.B., Jr., and Salas, E. (1990). *A comparison of two types of training interventions on team communication performance.* Paper presented at the Annual Meeting of the Human Factors Society, Orlando, FL.

Levis, A.H. (1989). Generation of architectures for distributed intelligence systems. *Proceedings of the 1989 IEEE International Conference on Control and Applications.* Jerusalem, Israel: IEEE Press.

McIntyre, R.M., Morgan, B.B., Jr., Salas, E., and Glickman, A.S. (1988). Teamwork from team training: New evidence for the development of teamwork skills during operational training. *Proceedings of the 10th Interservice / Industry Training Systems Conference.* (pp. 21-27). Orlando, FL: National Security Industry Association.

McMillan, G., Beevis, D., Salas, E., Strub, M.H., Sutton, R., and Von Breeda, A. (Eds.). (1989). *Application of human performance models for system design.* New York: Plenum Press.

Meichenbaum, D. (1985). Stress inoculation training. In Goldstein, A.P., Krasner, L. and Garfield, S.L. (Eds.), *Psychology practitioner guidebooks.* New York: Pergamon Press.

Morgan, B.B., Jr., Glickman, A.S., and Woodard, E.A. (1986). *Measurement of team behaviors in a Navy training environment.* (Report No. TR-86-014). Norfolk, VA: Old Dominion University, Center for Applied Psychological Studies.

Noble, D. (1989). *Application of a theory of cognition to situation assessment.* Engineering Research Associates (Office of Naval Research contract N00014-84-0484), Vienna, VA.

Novaco, R., Cook, T., and Sarason, I. (1983). Military recruit training: An arena for stress-coping skills. In Meichenbaum, D. and Jaremko, M. (Eds.), *Stress reduction and prevention.* New York: Plenum.

Orasanu, J. (1990). *Shared mental models and crew performance.* (Tech. Rep. No. 46). Princeton, NJ: Princeton University.

Orasanu, J. and Salas, E. (in press). Team decision making i complex environments. In Klein, G., Calderwood, R. and Orasanu, J. (Eds.), *Decision making in action: Models and methods.* Norwood, NJ: Ablex.

Oser, R., McCallum, G.A., Salas, E., and Morgan, B.B., Jr. (1989). *Toward a definition of teamwork: An analysis of critical team behaviors.* (Report No. 89-004). Orlando, FL: Naval Training Systems Center.

Payne, J., and Bettman, J. (1988). Adaptive strategy selection in decision making. *Journal of Experimental Psychology: Learning, Memory and Cognition, 14* (3), 534-552.

Pennington, N., and Hastie, R. (in press). A theory of explanation-based decision making. In Klein, G.A., Orasanu, J. and Calderwood, R. (Eds.), *Decision making in action: Models and methods*. Norwood, NJ: Ablex.

Phelps, R.H., Pliske, R.M., and Mutter, S.A. (1987). Improving decision making: A cognitive approach. In Zeider (Ed.), *Human productivity enhancement*. New York: Praeger.

Reisig, W. (1985). *Petri nets*. New York: Springer-Verlag.

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

Sage, A.P. (1981). Behavioral and organizational considerations in design of information systems and processes for planning and decision support. *IEEE Transaction on Systems, Man and Cybernetics, SMC-11,* 640-678.

Salas, E., Blaiwes, A., Reynolds, R.E., Glickman, A.S., and Morgan, B.B., Jr. (1985, Nov.). Teamwork from team training: New directions. *Proceedings of the 7th Interservice/Industry Training Equipment Conference and Exhibition*, 400-406.

Shriffrin, R.M., and Schneider, W. (1977). Controlled and automatic human information processing: II. Perceptual learning, automatic attending, and a general theory. *Psychological Review, 84,* 127-190.

Simon, H. (1957). *Models of man: Social and rational.* New York: Wiley.

Stout, R., Cannon-Bowers, J.A., Salas, E. and Morgan, B.B. (1990). Does crew coordination behavior impact performance? *Proceedings of the 34th annual meeting of the Human Factors Society.* (pp. 1382-1386). Santa Monica, CA: Human Factors Society.

Tolcott, M.A. (1991). *Understanding and aiding military decisions.* Paper presented at the 27th International Applied Military Psychology Symposium, Stockholm, Sweden.

Tolcott, M.A. and Holt, V.E. (1987). *Impact and potential of decision research on decision aiding.* Technical Report. Washington, DC: American Psychological Association.

Veldhuyzen, W., and Stassen, H.G. (1977). The internal model concept: An application to modeling human control of large ships. *Human Factors, 19,* 367-380.

Wickens, C.D. (1984). *Engineering psychology and human performance.* Columbus, OH: Charles E. Merrill.

Williams, M.D., Hollan, J.D., and Stevens, A.L. (1983). Human reasoning about simple physical systems. In Gentner, D. and Stevens, A.L. (Eds.), *Mental models.* Hillsdale, NJ: Erlbaum.

Zakay, D., and Woller, S. (1984). Time pressure, training, and decision effectiveness. *Ergonomics, 27,* 273-284.

REPORT DOCUMENTATION PAGE

Form Approved
OMB No 0704-0188

1. AGENCY USE ONLY (*Leave blank*)

2. REPORT DATE
August 1992

3. REPORT TYPE AND DATES COVERED
Applied Military Psychology

4. TITLE AND SUBTITLE
The 27th International Applied Military Psychology Symposium: A Focus on Decision Making Research

5. FUNDING NUMBERS

6. AUTHOR(S)
Stanley C. Collyer

7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES)
Office of Naval Research
European Office
PSC 802 Box 39
FPO AE 09499-0700

8. PERFORMING ORGANIZATION REPORT NUMBER
92-4-W

9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES)

10. SPONSORING/MONITORING AGENCY REPORT NUMBER
ONREUR
Report 92-4-W

11. SUPPLEMENTARY NOTES

12a. DISTRIBUTION/AVAILABILITY STATEMENT
This report is unclassified; distribution is unlimited.

12b. DISTRIBUTION CODE

13. ABSTRACT (*Maximum 200 words*)

The first International Applied Military Psychology Symposium (IAMPS) was held in 1963. Since then it has evolved into an important forum for information exchange among European and North American military psychologists. A wide array of topics are typically covered. These include personnel selection and classification, manpower planning, leadership, stress, training, simulation, and human factors.

14. SUBJECT TERMS
Military Psychology
sporadic processes
constricted processes
decision-making
fluid processes

15. NUMBER OF PAGES
74

16. PRICE CODE

17. SECURITY CLASSIFICATION OF REPORT	18. SECURITY CLASSIFICATION OF THIS PAGE	19. SECURITY CLASSIFICATION OF ABSTRACT	20. LIMITATION OF ABSTRACT
UNCLASSIFIED	UNCLASSIFIED	UNCLASSIFIED	UL