(1)

Quarterly Technical Report PQTR-1046-78-9
Contract MDA903-77-C-0184
ARPA Order No. 3344
Report Date September 1, 1978

AN INTERACTIVE COMPUTER AIDING SYSTEM FOR GROUP DECISION MAKING

ANTONIO LEAL
RANDALL STEEB
STEVEN JOHNSTON
STEVEN LEVIN

Prepared for:

Cybernetics Technology Office
DEFENSE ADVANCED RESEARCH PROJECTS AGENCY
1400 Wilson Boulevard
Arlington, Virginia 22209

PERCEPTRONICS

6271 VARIEL AVENUE • WOODLAND HILLS • CALIFORNIA 91367 • PHONE (213) 884-7470

NOTES

Quarterly Technical Report PQTR-1046-78-9
Contract MDA903-77-C-0184
ARPA Order No. 3344
Report Date September 1, 1978

AN INTERACTIVE COMPUTER AIDING SYSTEM FOR GROUP DECISION MAKING

ANTONIO LEAL
RANDALL STEEB
STEVEN JOHNSTON
STEVEN LEVIN

Prepared for:

Cybernetics Technology Office
DEFENSE ADVANCED RESEARCH PROJECTS AGENCY
1400 Wilson Boulevard
Arlington, Virginia 22209

PERCEPTRONICS
6271 VARIEL AVENUE • WOODLAND HILLS • CALIFORNIA 91367 • PHONE (213) 884-7470

UNCLASSIFIED

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

REPORT DOCUMENTATION PAGE		READ INSTRUCTIONS BEFORE COMPLETING FORM
1. REPORT NUMBER PQTR-1046-78-9	2. GOVT ACCESSION NO. AD-A134 939	3. RECIPIENT'S CATALOG NUMBER
4. TITLE (and Subtitle) AN INTERACTIVE COMPUTER AIDING SYSTEM FOR GROUP DECISION MAKING		5. TYPE OF REPORT & PERIOD COVERED TECHNICAL REPORT
		6. PERFORMING ORG. REPORT NUMBER
7. AUTHOR(s) Antonio Leal Steven Johnston Randall Steeb Steven Levin		8. CONTRACT OR GRANT NUMBER(s) MDA903-77-C-0184
9. PERFORMING ORGANIZATION NAME AND ADDRESS Perceptronics, Inc. 6271 Variel Avenue Woodland Hills, California 91367		10. PROGRAM ELEMENT, PROJECT, TASK AREA & WORK UNIT NUMBERS ARPA Order No. 3344
11. CONTROLLING OFFICE NAME AND ADDRESS DARPA - Cybernetics Technology Office 1400 Wilson Boulevard Arlington, VA		12. REPORT DATE September 1978
		13. NUMBER OF PAGES
14. MONITORING AGENCY NAME & ADDRESS(if different from Controlling Office)		15. SECURITY CLASS. (of this report) UNCLASSIFIED
		15a. DECLASSIFICATION DOWNGRADING SCHEDULE

16. DISTRIBUTION STATEMENT (of this Report)

Approved for Public Release; Distribution Unlimited

17. DISTRIBUTION STATEMENT (of the abstract entered in Block 20, if different from Report)

18. SUPPLEMENTARY NOTES

Approved by: Antonio Leal

19. KEY WORDS (Continue on reverse side if necessary and identify by block number)

Group Decision Aiding
Decision Trees
Utility Conflict Resolution
Utility Aggregation
Sensitivity Analysis

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

This report describes a program centered on the demonstration of an interactive computer aid for group decision making. The report includes: 1) descriptions of technical improvements implemented on the system, 2) details of system improvements now in progress, and 3) an analysis of tests made to determine the applicability of the set of CACI-developed system evaluation measures.

DD FORM 1 JAN 73 1473 EDITION OF 1 NOV 65 IS OBSOLETE

UNCLASSIFIED

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

TABLE OF CONTENTS

	Page
1. SUMMARY	1-1
1.1 Report Period	1-1
1.2 Next Period	1-2
1.3 Program Milestones	1-2
2. PROGRAM OVERVIEW	2-1
2.1 Statement of Problem	2-1
2.2 Rationale	2-1
2.3 Objectives	2-2
3. SYSTEM IMPROVEMENTS	3-1
3.1 General	3-1
3.2 Implemented Improvements	3-2
3.3 Improvements in Progress	3-7
4. TESTS OF CACI EVALUATION INSTRUMENT	4-1
5. REFERENCES	5-1
APPENDIX A	A-1

1. SUMMARY

1.1 Report Period

The sixth quarter of contract activity involved: (1) implementation of changes identified during initial test and evaluation experiments, (2) specification of additional technical improvements, and (3) tests of procedures for evaluation of the group aid. The following tasks were completed during the report period:

(1) A series of changes were completed following test and evaluation experiments and system demonstrationa included were: (a) algorithms for sensitivity analysis override and tree traversing, (b) procedures for scale definition for attributes and for overall values, and (c) provision for individualized attribute weighting.

(2) Additional technical improvements to the group decision aid were begun in parallel with those described above. Among the changes are: (a) stored descriptions of major alternatives, events, and attributes, (b) interactive sensitivity analysis, displayed graphically for both probabilities and values, and (c) new conflict resolution algorithm, based on individual differences at the root node.

(3) Experiments were performed to determine the operational applicability of the system evaluation measures and methods developed by CACI. Both aided and unaided group sessions were analyzed using the evaluation instrument. Some modifications to the procedures were made, and a set of guidelines for future application were developed.

1.2 Next Period

The contract period during the next quarter will primarily concentrate on the finalization and testing of the system improvements and extensions. The specific items of work for the next period include:

(1) Implement and test system improvements and extensions.

(2) Transfer system to NADC for demonstration.

(3) Prepare for preliminary experimental evaluation of system; perform evaluation.

1.3 Program Milestones

The milestone chart for the contract program is shown in Table 1-1, with the report period illustrated as the checkered portion.

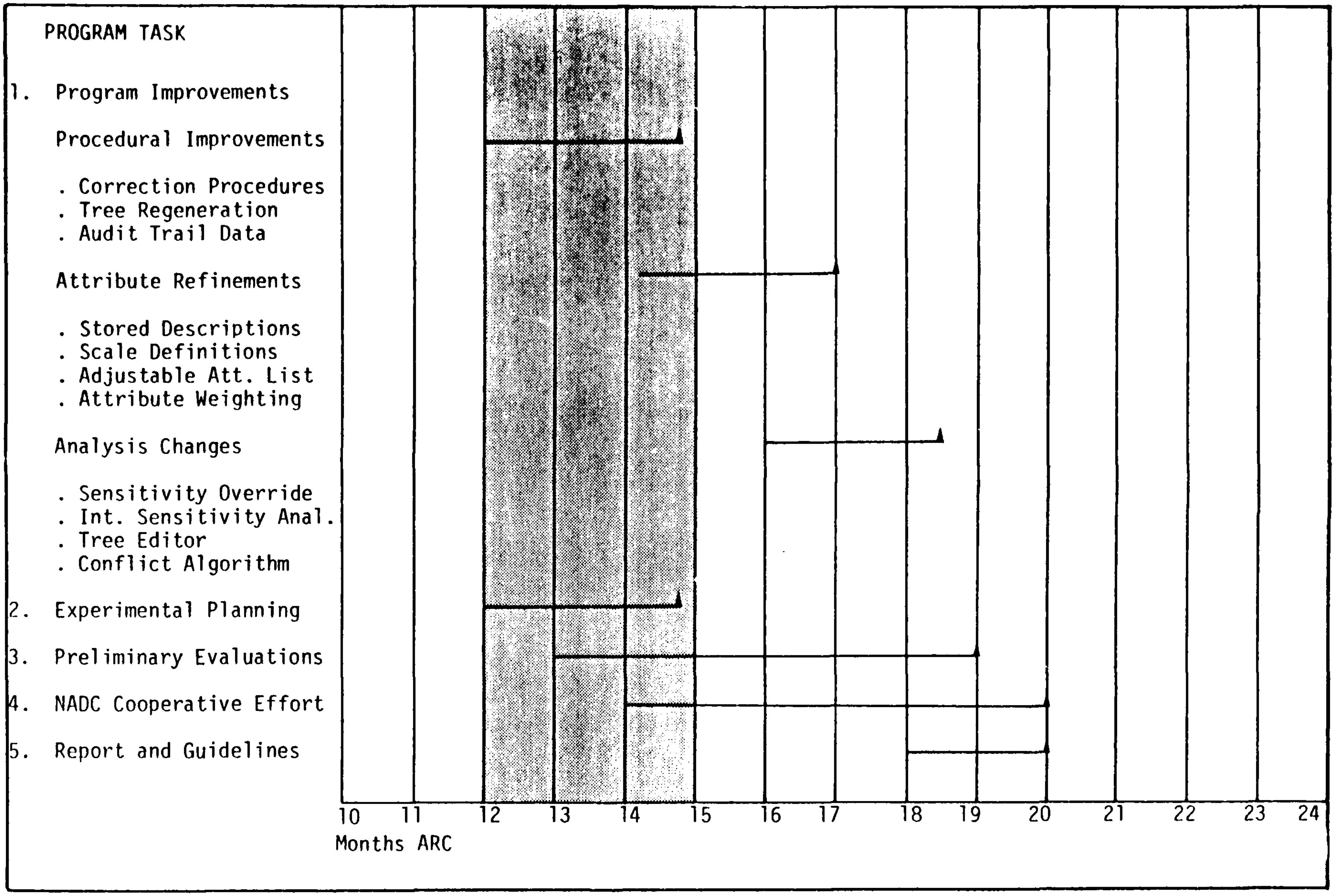

TABLE 1-1. PROGRAM MILESTONES

2. PROGRAM OVERVIEW

2.1 Statement of Problem

Constant escalation in weapons cost and effectiveness, as well as the increasing complexity of international relations, makes military decision making more critical today than ever before. In today's military environment, most upper-level decisions are made by committees and staff groups. Typically, such groups contain experts from several speciality areas, who bring to the decision environment disparate sets of values. Decision time is usually limited, the decision making procedure is relatively unstructured, and intragroup conflicts arise on a broad variety of issues. Consequently the group is unable to consider the maximum set of alternatives, conflicts are not resolved in an optimum manner, and the resultant decision is rarely up to the aggregate potential of the group membership.

2.2 Rationale

Decision analysis offers a promising approach to solving these problems. The analytical procedure of building a decision tree formalizes the decision process, and permits incorporation of individual values (utilities) into the selection of alternative courses of action (Hays, O'Connor, and Peterson, 1975). However, decision analysis as it is usually practiced, is a highly personal and time-consuming process. Decision analysts are often called upon to assist in the solution of problems ranging over a large variety of domains. In most cases, the decision analysts know far less about the problem domain than do their clients. Thus their contributions are confined primarily to the phases of formalization and optimization. While optimization is usually computer assisted, the formalization phase invariably has been accomplished manually, using lengthy interviews of persons more familiar with the problem area. This approach is generally incompatible with the conditions of command group decision making.

Accordingly, it would be highly worthwhile to automate the formalization phase, using an interactive computer system to interrogate the group members and to construct a decision tree based on their responses. The purpose of the research undertaken here is to develop and evaluate the means by which such an interactive aid could be used to improve group decision making.

2.3 Objectives

The goal of the research program addressed in this progress report is to develop an automated decision tree elicitation system using on-line sensitivity analysis with direct real-time group feedback and evaluate its effectiveness in aiding group decision making.

The specific objectives of the current program include the following:

(1) Develop computer programs for efficient, comprehensive, elicitation of decision trees from a decision making group.

(2) Develop computer programs for identifying structural and numerical differences among the contributions of individual group members, for merging these contributions and for resolving the points of conflict.

(3) Develop effective means for displaying to the group the results of the elicitation procedures and conflict analyses.

(4) Integrate the various programs and techniques into a complete aiding system which can be readily transferred to other test environments.

(5) Experimentally test the Group Decision Aid, using a variety of representative military decision problems, to demonstrate its advantages under realistic conditions of use.

(6) On the basis of the developmental effort and the experimental results, establish guidelines and recommendations for future military applications of the group decision aiding methodology.

3. SYSTEM IMPROVEMENTS

3.1 General

The initial system tests, described in our recent technical report (Leal, et al, 1978), were highly successful. The initial development goals of the system were realized, in that experienced groups were able to work effectively with the aid toward the solution of a realistic and representative decision problems. At the same time, a number of system areas were identified in which existing features could be modified, or new features added, in order to improve overall system performance and power. Suggestions for candidate improvements came (1) from our review of original system goals; (2) from our review of related developments in computerized decision aiding, such as the IBM 5100 aiding packages produced for ARPA/CTO by DDI; (3) from suggestions made by knowledgeable visitors to whom the system was demonstrated, such as Dr. Clinton Kelly, of DDI; Dr. Martin Tolcott of ONR; Dr. Ward Edwards of USC, and others; and (4) from suggestions made by our ARPA monitors and other ARPA personnel.

The improvements selected for implementation during the FY78 period, i.e., by September 30, 1978, include modifications for tree traversing, scale definition, attribute weighting, stored descriptions, interactive sensitivity analysis, and new conflict identification algorithms. The first three of these, tree traversing, scale definitions, and attribute weighting, have now been implemented. Section 3.2 describes these changes. Plans for the remaining changes, now in progress, are detailed in Section 3.3.

3.2 Implemented Improvements

The following changes to the group decision aiding system have been completed. These modifications enhance the decision theoretic fidelity of the aid and are particularly important for the NADC cooperative effort.

Tree Traverser. It is unrealistic to expect the group to be able to develop a decision tree without error the first time. Alternatives will be forgotten; events will be found too improbable to be included; formally closed branches will be reopened, etc. A tree restructuring capability has been implemented which permits groups to alter previously build protions of the tree at any time. Movement to the decision node desired is shown by a cursor under the control of the intermediator.

Scale Definitions. It is critical to define clearly the endpoints and properties of the utility scales. The scales of both the attributes and the action/event nodes must be defined consistently. Each attribute is scaled by defining the worst possible level to be zero and the best possible to be 100. Shortened labels are assigned at each of these endpoints, such as in the example below.

$100 million Loss	$20 million Loss
0	100

Suggestions have been made for a reversible scale, on which the zero point will be the best situation, e.g., zero hostages killed, zero dollars lost, etc. However, this would make translations from attribute scales to overall value scales difficult. The use of well-defined endpoints is expected to solve this problem.

The overall utilities of actions or events are similarly defined by labeling a single, common scale, established early in the analysis. Again, the zero point corresponds to the worst possible state, while the 100 point denotes the best possible state. At the same time, consistency is maintained by having the upper and lower endpoints correspond to all zeros and all 100s on the individual attributes, respectively. Finally, for flexibility, changes can be made to the labels at any time during the task.

Adjustable Attribute List. The construction of the attribute list is the first task the group must perform. However, it is very difficult to foresee the applicability of the attributes to particular future alternatives or outcomes. An adjustable attribute list has been included to allow new attributes to be added and to allow irrelevant ones to be deleted.

For the most part, the attribute set should be invariant with respect to the alternatives. This is because the attribute membership derives from the set of objectives characterizing the problem, and not from the choices present. However, development of the decision tree may bring to light additional considerations which should be included as attributes. For ease of comparability across actions, such changes may have to be incorporated in previously evaluated nodes.

Attribute Weighting. It is unlikely that the attributes will all be equally important with respect to the particular alternative in conflict. A number of researchers (Newman, Seaver, and Edwards, 1976; McClelland, 1978) have shown that differential weighting (distinct weights for each attribute) is important in most real-world situations. The evaluation function will then take the following form:

$$U_j = \sum_i w_i A_{ij}$$

where U_j is the overall utility of outcome J

w_i is the importance weight of attribute i

A_{ij} is the level of attribute i occurring with outcome j

The weights w_i define the policy of the decision maker with regard to the different dimensions of outcome. This policy should be invariant with respect to the possible alternatives. The attribute levels, on the other hand, characterize the choices or outcomes along the dimensions -- costs, tactical gains, political impact, etc. As such, these levels vary with each choice and must be estimated for each point in conflict.

The importance weights will be elicited from each participant after definition of the set of attributes. The elicitation process is patterned after that used by DDI (Selvidge, 1976) and by SSRI (Gardiner and Edwards, 1975). In short, comparisons will be made of the importance of swings across the range of each attribute. One of the attributes will arbitrarily be given a weight of ten points. Each remaining attribute will be compared to the first attribute by estimating the importance of a change from the lowest to highest level compared to a change from lowest to highest on the first attribute. For example, if a swing across the range of the second attribute is twice as important as a swing across the range of the first attribute, a weight of 20 will be assigned. Once all of the attributes are thus weighted, a normalization is made: Each attribute is set equal to its raw weight divided by the sum of raw weights and then multiplied by 100. This results in a summation of all weights to 100 points.

The frames used during the attribute weighting process are shown in Figures 3-1 and 3-2. Figure 3-1 depicts one of the attribute comparisons

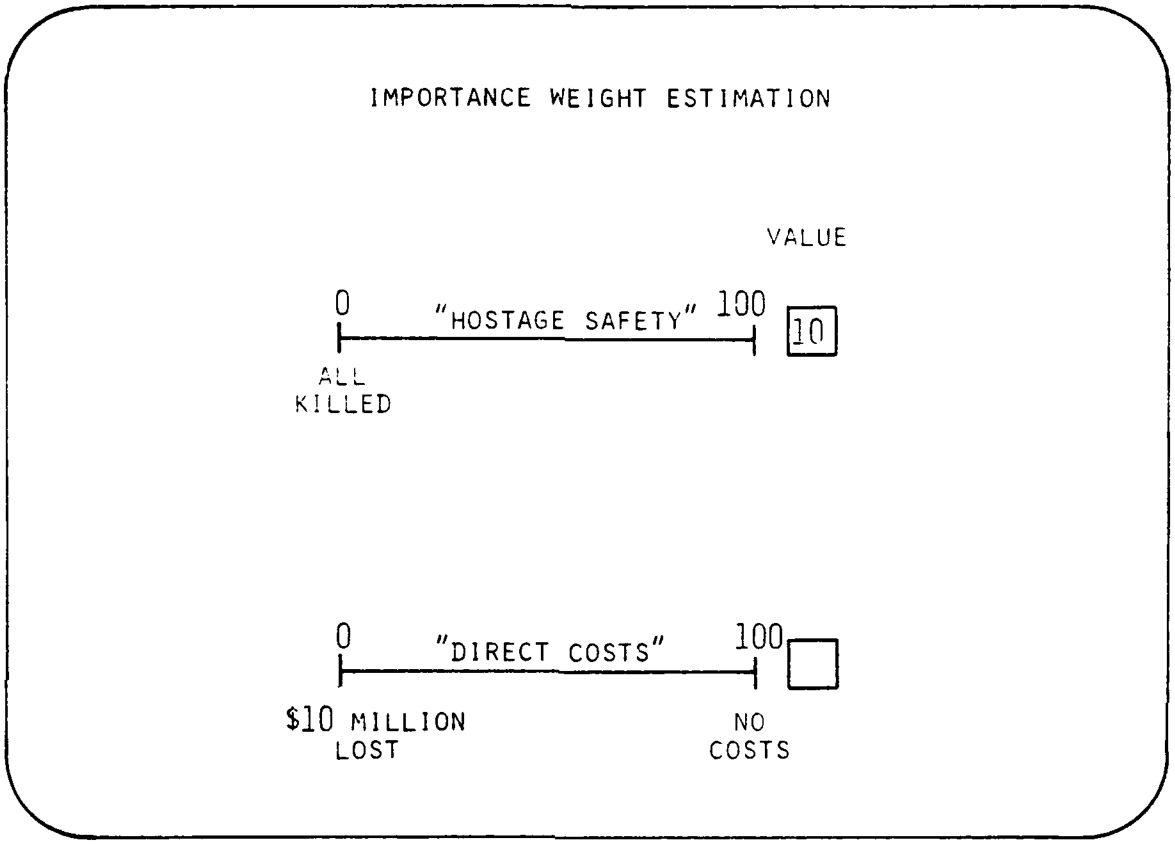

FIGURE 3-1. ATTRIBUTE WEIGHTING FRAME

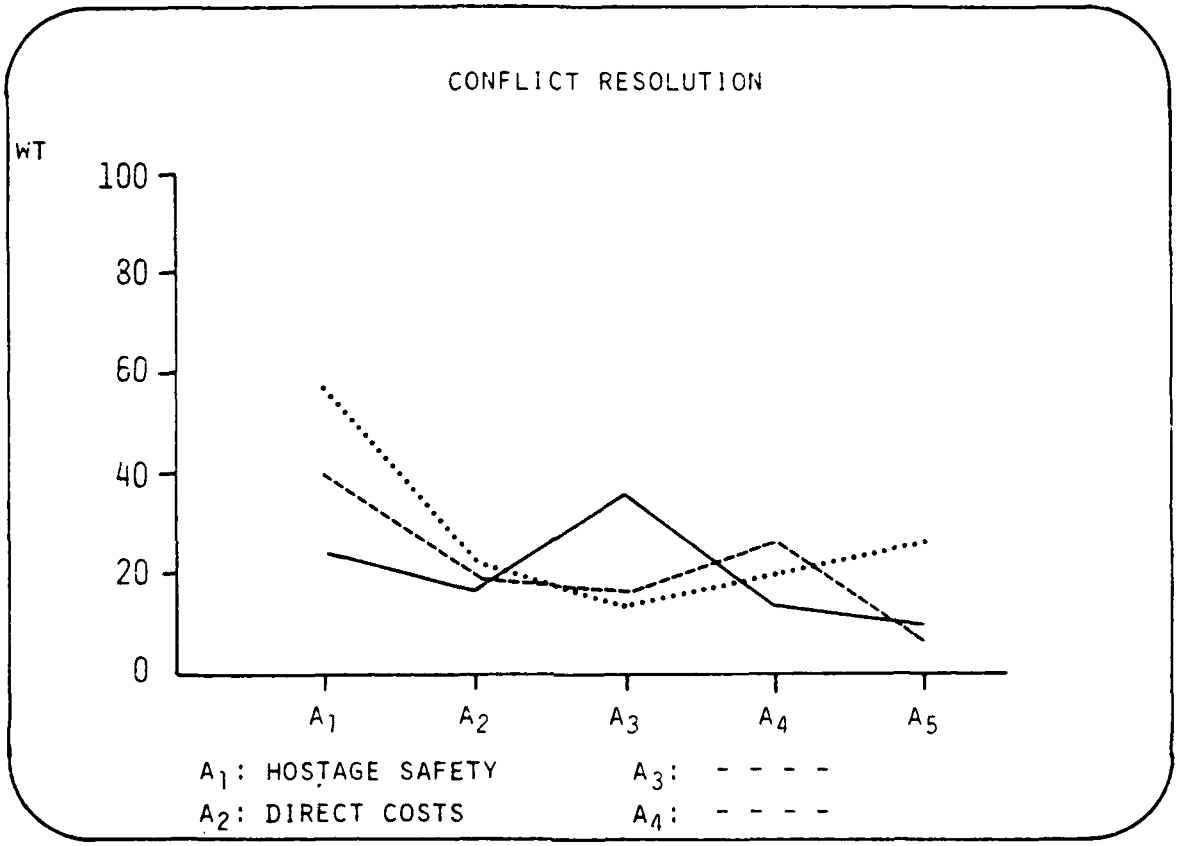

FIGURE 3-2. WEIGHTING CONFLICT RESOLUTION FRAME

described above. The upper scale in the figure is used as the baseline, and is assigned a value of 10 (shown at the right of the frame). Each of the remaining attributes is compared in turn to this attribute.

If there is a conflict in estimation after the normalization process, the frame shown in Figure 3-2 is displayed. This is a graphical display of the normalized individual estimates. Discussion and re-estimation using the scale comparison technique then follows. The arithmetic mean of the estimates is used in all later calculations, although the final individual weight vectors are stored for possible later tests of conflict.

3.3 Improvements in Progress

The following changes to the group system are now in progress. These improvements are schedules to be implemented by October 1, 1978.

Interactive Sensitivity Analysis. One of the most important functions of the group decision aid will be to provide information about individual problem elements through sensitivity analysis. Sensitivity analysis tells how the relative attractiveness of the various actions change as each single aspect of the decision is varied. The parameters planned to be varied are the overall node values, the attribute levels, the attribute weights, and the event probabilities. Each of these will be described below.

A graphic format will be used for displaying the sensitivity analysis. Figure 3-3 illustrates an example display using as a variable the overall node value. The figure is a plot of the utility (the expected value) of actions 1, 2, and 3, the root actions, as the value of the node in question is varied across its possible range. The sensitivity shows

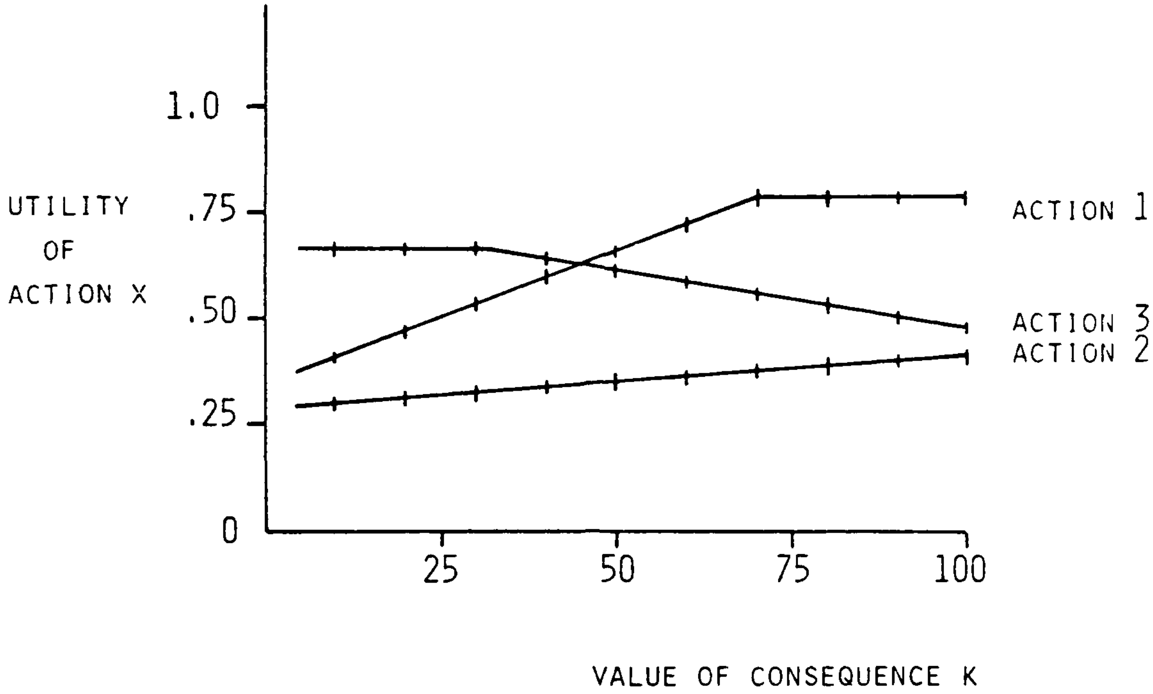

FIGURE 3-3. SENSITIVITY ANALYSIS REGARDING VALUE OF CONSEQUENCE K

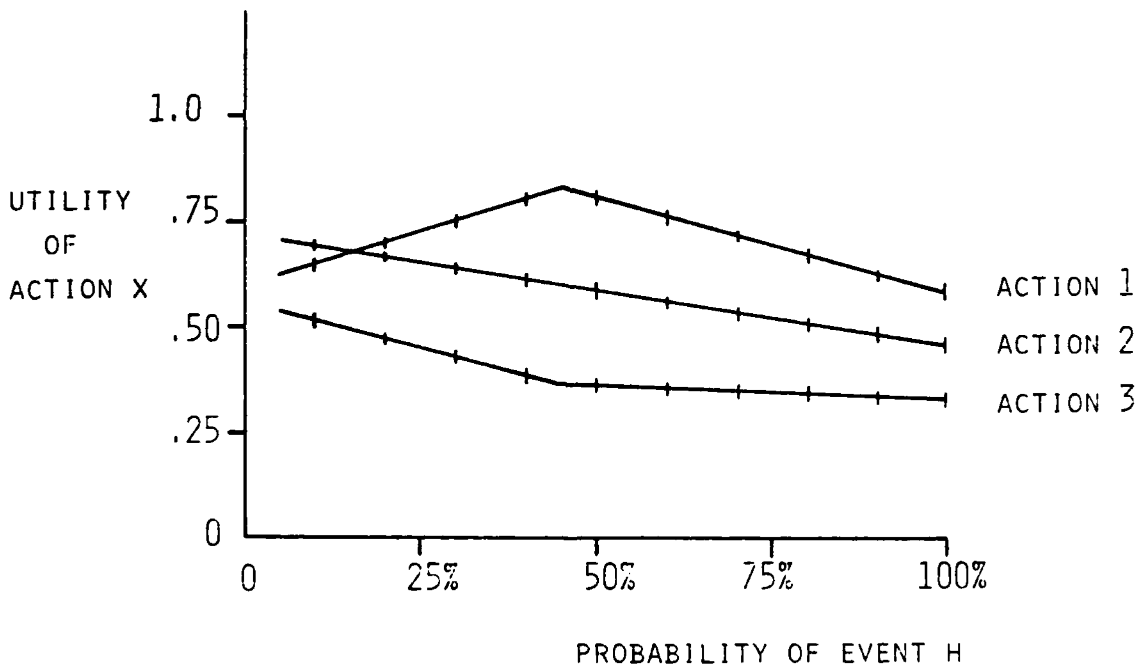

FIGURE 3-4. SENSITIVITY ANALYSIS REGARDING PROBABILITY OF EVENT H

information about crossover points, values at which the favored action changes. The example in Figure 3-3 shows action 3 favored until the node value goes above 50, after which action 1 becomes favored. For computational simplicity, samples of the values can be taken at 10 point intervals. This should provide sufficient information for making the plot.

Sensitivity analysis for the attribute levels and weights are slightly more complicated. For attribute level, the plot would be the same form as that of Figure 3-3, with attribute A_i defining the abscissa rather than value K. Sensitivity analysis for the attribute weights, on the other hand, would substitute the attribute weight w_i for value K. Both the attribute level and weight analyses require a complete folding back of the tree.

Sensitivity analysis with respect to probability of events is again similar. Figure 3-4 shows how a probability of event H can be varied across a 0 to 100% chance of occurrence to determine its effect on the values of the various actions. If more than two events are possible at the node in question, the ratio of likelihoods between the remaining events will be kept invarient. This form of sensitivity analysis is especially useful regarding value of information. Information that changes the probability distribution sufficient to justify its cost can be identified.

Each of the above sensitivity analyses will be available to the intermediator. Requests for a specific analysis will be made by the participants.

Stored Descriptions. It is sometimes difficult to capture complex alternatives or outcome situations with a few words (currently ten letters for tree descriptions). A small amount of time should be used by the group to develop detailed definitions and descriptions of each major alternative, event, and attribute. The descriptions may be displayed at any time for reference. The description may include such items as the actor, the action, the recipient, the situation, and the time frame. An example of an action description follows:

> Display Strength: Deploy carriers immediately to Mandero City and await response.

<u>New Conflict Resolution Algorithm</u>. The current conflict identification and resolution algorithm is somewhat arbitrary in the sense that there is only an indirect relationship between differences in values and differences in preferred actions. For example, it is possible to have large differences in opinion as to the value of an event, but agreement on the initial decision (at the root of the tree). A new algorithm will identify conflicts only as those differences of values that cause a difference in the initial decision. The sensitivity analysis program will test each participant's values, one at a time. If a conflict is present, the node just opened will be decomposed into attributes. The previous conflict algorithm concerning differences in attribute levels will then be used. After recombination, the overall values will again be tested for conflicts at the root decision. If a conflict is still present, the remaining nodes will, one by one, be decomposed into attributes until either the root decision conflict is resolved or no further discussion is desired.

Probability conflicts will be tested in a parallel fashion, by averaging the value estimates and determining if the probability differences result in choice differences at the root node.

4. TESTS OF CACI EVALUATION INSTRUMENT

CACI, operating as subcontractor to Perceptronics, produced in June, 1978, a comprehensive set of evaluation measures to test the group decision aiding system (CACI, 1978). In July, 1978, Perceptronics and CACI arranged to undertake "shakedown" runs on these measures by scheduling and analyzing several group sessions in a one-week period. Dr. Bert Spector of CACI assisted in these tests and performed much of the analysis. One of the two sessions involved use of the group decision aid, while the second group of three subjects was unstructured. Since the objective was one of testing the measurement instrument, college students were used, and no attempt was made to balance for age or background. Both groups dealt with the counter-terrorist scenario. Stereo audio recording was used to augment the audit trail for analysis of the sessions.

For the most part, the 23 measures of the evaluation instrument were found to be satisfactory for characterizing the aided session. About half of the measures were also found to be applicable to the unaided session analysis. The conflict resolution and utility loss measures were not applicable because values or probabilities were not estimated during the non-aided session. A memo describing modifications made to the measures, along with guidelines for application, was prepared by Dr. Spector. This memo "Applying Evaluation Methodologies to the Group Decision Aid," is attached as Appendix A.

While the sessions were not intended for comparison of conditions, some behavioral observations were made. The findings were interesting enough to warrant a replication of the unaided condition with a second group. This second group received additional debriefing to elicit specific values and probabilities.

All three groups, two unaided and one aided, used almost the full three hours to discuss options and come up with a course of action. The unaided groups tended to settle in on a plan early in the analysis and modify it to reflect new considerations and contingencies. One of the unaided groups, as a result, did not consider the immediate attack option at all, and simply expanded the stall and attack alternative. The other unaided group split into two factions, one supporting immediate attack (2 members) and the other supporting a stall and attack plan (1 member). These factions followed the classic group stratagem of finding points to support their own position. After approximately two hours of such conflict, the person favoring a stall tactic capitulated, although many of that member's considerations were incorporated in the final plan. All in all, the unaided groups were able to come up with a fairly rich analysis of the problem, mentioning many actions, events, and contingencies (in one case, producing a richer set than the aided group). This appeared to be because they gave very little attention to ranking or rating the alternatives, estimating probabilities, or considering attributes. The objectives of recovering the bombs and saving the hostages were often mentioned, but the set of considerations was never consistent.

The final recommendations produced by the unaided groups were quite detailed regarding the final plan. However, no secondary options and few rationalizations for the choice were provided (although these were requested). It appears that the unaided groups would have had a difficult time of communicating the rationale for their plan or changing the plan as new circumstances arose. The aided group produced a less detailed but much more systematically analyzed and communicable course of action.

Some recommendations for further work appear necessary. Several of the system changes planned for the next few months, notably the tree traversing and new conflict resolution procedures, appear indispensible.

Also, experienced decision makers familiar with the problem area should be used. Such individuals would be able to formulate courses of action and estimate consequences much more effectively than a scenario-trained group. Finally, a situation estimate scenario with established "ground truth" would be instrumental for providing objective analyses.

5. REFERENCES

Gardiner, P.C. and Edwards, W. Public Values: Multi-Attribute Utility Measurement for Social Decision Making. In Schwartz & Kaplan (Eds.), Human Judgment & Decision Processes: Formal and Mathematical Approaches. New York: Academic Press, Inc., 1975.

Hays, M.L., O'Connor, M.F., and Peterson, C.R. An Application of Multi-Attribute Utility Theory: Design-to-Cost Evaluation of the U.S. Navy's Electronic Warfare System. Decisions and Designs, Inc., (McLean, VA), Technical Report DT/TR75-3, October 1975.

Leal, A., Levin, S., Johnston, S., Agmon, M., and Weltman, G. An Interactive Computer Aiding System for Group Decision Making. Perceptronics Inc., (Woodland Hills, CA), Technical Report PQTR-1046-78-2, Contract MDA903-77-C-0184, ARPA Order No. 3344, Defense Advanced Research Projects Agency, February 1978.

McClelland, G.H. Equal Versus Differential Weighting for Multi-Attribute Decisions: There are no Free Lunches. University of Colorado Institute of Behavioral Sciences, Technical Report No. 207, 1978.

Newman, J., Seaver, D.A., and Edwards, W. Unit Versus Differential Schemes for Decision Making. Social Science Research Institute (Los Angeles, CA), Technical Report 001597-2-T, July 1976.

Selvidge, J. Rapid Screening of Decision Options. Decisions and Designs, Inc., (McLean, VA), Technical Report TR76-12, 1976.

APPENDIX A

APPLYING EVALUATION METHODOLOGIES
TO THE GROUP DECISION-MAKING AID:
A CACI TECHNICAL MEMORANDUM

INTRODUCTION

This memorandum documents the procedures, findings, and recommendations that resulted from several experiments in which evaluation measures and methods developed by CACI (1978a) were implemented. The system being evaluated was the Perceptronics (1978) group decision-making aid. The first section discusses the operational applicability of the evaluation measures to both experimental and control treatments and describes the further development of some of the measures to improve their comparability and monitoring ease. Second, procedures that should be followed in future evaluation experiments involving the group aiding system are detailed. Finally, recommendations are developed to improve the interpretability of future evaluation experiments.

OPERATIONAL EVALUATION MEASURES

Evaluation Exercises

On 18-19 July 1978, two experimental tests of the evaluation measures were conducted at the Perceptronics test site. One test involved a three-person team attempting to resolve a crisis management scenario (CACI, 1977) with the assistance of the group aiding system. The other test, a control treatment, involved a different three-person team confronted with the identical scenario and tasking, but given no aid or prompting whatsoever. Hence, the control group was free to develop its own procedures, rules, and methods for resolving the crisis scenario.

This free form control treatment, more than any other type of control group, is probably closest to ad hoc crisis management task forces that are formed in government to deal with real crisis situations. Any decision-making structure or formalization of procedures that evolves during task force operations is generated internally and is not imposed by external sources. On the other hand, the experimental treatment, involving the use of the group decision-making aid, imposes the structure and processes of format decision analysis techniques on the decision-making team.

The evaluation measures were able to monitor sensitive assessment criteria for both the experimental and control treatments. Comparison of identical measures for each treatment provided a systematic baseline against which the performance differential produced by using the group aiding system could be judged. The current set of evaluation exercises provides ample evidence of the applicability and utility of the evaluation measures developed by CACI to assess the Perceptronics group aiding system. The procedures and recommendations resulting from this evaluation comparison are described in this memorandum. However, due to the small number of cases (N=2), sampling deficiencies, and the transitional state of the group aiding system itself, the evaluation scores calculated from the experimental data do not validly represent the operational utility of the system. Hence, these figures will not be presented in this memorandum.

Applicability of Evaluation Measures to Control Treatments

Most of the evaluation measures developed by CACI to monitor team performance with the group aiding system can be employed with little or no modification to monitor team performance in control settings as well. Due to the absence of any structure imposed by the experimenter in the control treatment, it is difficult to predict procedures that the control group will employ and, thus, to derive a set of evaluation measures that is jointly applicable to experimental and control treatments under all conceivable circumstances. However, there are likely to be several important commonalities between the performance of experimental and control groups that enable equivalent measurement. First, tasking for both types of groups in the current scenario requires that teams produce a recommended course of action at the conclusion of their sessions. Second, in their deliberations team members in either type of group are likely to generate various alternative actions and evaluate the extent to which these actions are likely to achieve team-valued objectives. These common performance elements enable compatible measurement of the aided (experimental) and unaided (control) groups using the same evaluation measures developed by CACI (1978a).

Table 1 presents the evaluation measures from CACI's (1978a) report that are compatible with data derived from unaided (control) groups. Most of the measures are directly compatible; that is, data that are likely to be obtained from control treatments will code easily into the categories required by the evaluation indices. A few of the evaluation measures will require the data obtained from unaided groups to be transformed into decision tree format to enable comparable measurement. Only four evaluation measures are likely to be incompatible in monitoring the control treatments. Each measure in the last category involves indexing of value assignments, which are not likely to be generated by control groups unless they employ formal decision analysis procedures.

FURTHER DEVELOPMENT OF EVALUATION MEASURES

Application of the evaluation measures in the evaluation exercises helped to refine and further develop some of the indices to make them more comparable in future experiments. Moreover, additional measures were suggested in analyzing the exercise data. These developments are discussed below.

Measures 2, 3, 4, and 6. Time Measures

Each measure can be expressed as a percentage of the total time expended to complete the decision-making session (Measure 1). Together, these four measures will equal 100 percent.

TABLE 1

Evaluation Measures Compatible With Unaided (Control) Groups

Measure	Direct Correspondence	Requires Decision Tree Transformation	Not Applicable
Task Analysis Measures			
1. Time expended to complete decision tree	x		
2. Time expended on information acquisition	x		
3. Time expended on objectives	x		
4. Time expended on alternatives generation	x		
5. Time expended on each alternative	x		
6. Time expended on alternative evaluation	x		
Tree Complexity Measures			
7. Total number of alternatives examined	x		
8. Total alternatives in final tree		x	
9. Percent nonproductive alternative generation	x		
10. Percent alternatives not in final recommendation	x		
11. Tree scope		x	
12. Percent effort devoted to contingency planning		x	
Participation Measures			
13. Participation time per member	x		
14. Mean number of alternatives generated	x		
15. Mean number of initiatives	x		
16. Rate of supportive output	x		
17. Rate of contributive output	x		
18. Rate of nonsupportive output	x		
Conflict Resolution Measures			
19. Percent utility conflicts reduced (without Multi-Attribute Utility Model (MAUM))		x	
20. Percent attribute conflicts reduced			x
Cohesion Measure			
21. Coalition formation			x
Informal Leadership Measure			
22. Implied influence			x
Power Distribution Measure			
23. Percent persuasive statements initiated	x		
Decision Quality Measures			
Content quality coefficient	x		
Process quality coefficient	x		
Outcome quality coefficient	x		
Subjective Attitudinal Measures			
Post-session survey	x		

Measure 10. Percent Nonproductive Branching

For clarification purposes, this measure should be relabeled, "Percent Alternatives Not Included in the Final Recommendation."

Measures 11 and 12. Tree Scope and Contingency Planning

For clarification purposes, "tree scope" is defined as the total number of choice and event points in the tree. The "percent effort devoted to contingency planning" is defined as the total number of event points divided by the total number of choice and event points in the tree.

Measures 19 and 20. Conflict Resolution Measures

Measure 19 monitors the reduction of utility differences without using the Multi-Attribute Utility Model (MAUM) or averaging procedures and, as such, does not apply to the group aiding system as it is presently configured. This measure is intended to tap the degree to which conflicting group value assignments on an action or event alternative are resolved on subsequent revotes after group discussion. As presently formatted, this cannot be done with the group aid. If a conflict arises, the group can only choose either the MAUM or averaging procedures; it cannot choose to discuss the conflicting positions and then vote again on the alternative. The aid would be greatly improved if the choice to conduct revotes on conflicting value assignments immediately after discussion were made available.

Measure 20 deals with the average reduction of attribute differences. This index taps those utility votes on attributes within the MAUM procedure that continue to be in conflict and must be resolved through additional votes on the conflicting attributes. To make the results of this measure more comparable, the outcome of the current equation (see Page C-7 of CACI, 1978a) should be divided by the first element in that equation (dealing with prior utility differences), and then multiplied by 100. The resulting measure can be relabeled, "Percent Attribute Conflicts Reduced."

Measure 21. Coalition Formation Measure

The data elements included in calculating this measure deal exclusively with strategic utility votes, that is, votes where there are significant differences in opinion among team players. To make this measure more comparable, the quotient of the current equation (see Page C-8 of CACI, 1978a) should be divided by 200 (for three-person teams), subtracted from unity, and then multiplied by 100. The resulting index measures the degree to which there is utility agreement among all three team members on alternatives in which there was initial conflict.

Measure 22. Implied Influence Measure

To improve comparability, the outcome of the current equation (see Page C-9 of CACI, 1978a) should be subtracted from unity. The resulting score is interpreted as the percent averaged deviation between a player's and the group's weighted utility assignments for each extended node.

Content and Outcome Quality Measures

The scenario employed in the 18-19 July exercises is different from the scenario originally developed by CACI (1977). All facts relating to U.S. troops at Komsa were omitted, and data concerning the chances for success in an attack were added. Because these critical changes were not communicated to CACI, the content and outcome quality measures must be adjusted. The fifth ranked criterion for content quality, "provide supplies/assistance to U.S. troops at Komsa," and the seventh ranked criterion for outcome quality, "actions to provide assistance to U.S. troops at Komsa," must be omitted and the previous ranks reduced by one.

New Measures Developed

Measure 20A. Percent Utility Conflicts Reduced After MAUM Procedures. This measure uses identical calculations as Measure 20. However, the utility votes that are compared are

- Initial value assignments on alternatives in which there is conflict, and
- Final value assignments on alternatives resulting from successful MAUM procedures.

Measure 4A. Percent Time Expended Resolving Value Conflicts. This measure is based on the calculation of time spent in MAUM or revote procedures.

Measure 12A. Percent Contingency Planning Activity. This measure divides the number of event alternatives by the total number of alternatives in the tree and multiplies the quotient by 100.

Measure 18A. Control Group Procedures. Of necessity, this is a free form monitoring of the basic procedures or structures instituted by the control group. Data on how the control groups operated can provide important information to categorize and characterize control group performance. For example, the results of extensive evaluation experiments may suggest that teams using the group aid perform better than control groups, which choose autocratic and nonparticipative procedures, but worse than control groups that prefer participative brainstorming techniques.

EVALUATION PROCEDURES

The exercises conducted on 18-19 July can be viewed as preliminary tests of the evaluation methods and procedures. Thus, they can be used as a learning device to improve future experiments. Some of the salient lessons are discussed on the following pages.

Subject Selection

One of the most critical lessons to come out of the exercises concerns subject selection. Greater control must be administered in choosing subjects to participate in the experiments, controlling team composition, and ensuring equivalency between experimental (aided) and control (unaided) treatments.

First, it is necessary to select subjects that are capable of dealing with a complex scenario on a sophisticated level. The quality of group deliberations and the practicality of the final recommended solution depend on the sophistication of the participants in handling the subject matter in the scenario. More sophisticated subjects are likely to yield higher quality decisions. Typically, naive subjects confronted with a complex scenario do not have sufficient background to know what the available response actions are. Thus, their recommended solutions often appear to be extremely general and too vague for practical implementation.

In the case of the current international crisis scenario, subjects require the capacity for strategic and tactical thought, as well as some experience, background, or knowledge concerning available options in such crisis situations. In the short term, the best civilian subject pool for the current scenario would probably include undergraduate political science and military science (ROTC) students.

Second, it is essential to control for team composition. As discussed in CACI (1978b), if adequate controls over team composition are absent, alternative explanations for team performance and the quality of team solutions may be posed. That is, without sufficient controls, it may be possible to assert that superior performance was obtained due to factors such as prior experience in team decision-making or prior use of decision analysis techniques by team members, rather than attributing the superior performance to the use of the group decision-making aid. Obviously, the point is to avoid the impact of such rival hypotheses. This can be done by anticipating and controlling for them through subject selection and team formation criteria.

Table 2 presents nine rival hypotheses that might be anticipated using the current group aid and the current scenario. Each of these plausible rival hypotheses comes from the scholarly literature on group dynamics in decision-making teams. On the basis of these hypotheses, it is possible to identify the characteristics of the ideal subject. They include

- Limited or no decision analysis training,
- Limited or no team decision-making experience,
- Limited or no experience with computer-based systems,
- Limited or no crisis management experience,
- Limited or no familiarity with other team members,
- Moderate levels of aspiration, and
- Moderate to extensive background in political science, international relations, or military science.

A subject pool that is larger than the required number of subjects should be interviewed by phone prior to the experiments to identify those persons who best conform with the characteristics of an optimal subject. These persons should be selected, and less than optimal subjects should be rejected.

Third, there should be equivalency between the aided and unaided groups. Subjects should be scheduled to form balanced teams. That is, subjects should be matched and assigned to teams so that the age and sex differences among teams are minimized.

TABLE 2

Rival Hypotheses

Background Hypotheses

1. If subjects had extensive experience working in similar decision-making teams, their decisions and performance are likely to be improved despite use of the Perceptronics aid (Stein, 1975).

2. If subjects had previous extensive experience or instruction using decision analysis methods, their decisions and performance are likely to be improved despite use of the Perceptronics aid (Goldman, 1965).

3. If subjects had previous extensive experience using computer-based systems, their decisions and performance are likely to be improved despite use of the Perceptronics aid (CACI, 1976b).

4. If subjects had previous extensive experience dealing with similar crisis management problems, their decisions and performance are likely to be improved despite use of the Perceptronics aid (Lorge and Solomon, 1960).

5. If subjects are extremely familiar with their fellow team members, their decisions and performance are likely to be improved despite use of the Perceptronics aid (Greer, et al., 1954).

6. If subjects have high levels of aspiration, their decisions and performance are likely to be improved despite use of the Perceptronics aid (Lorge and Solomon, 1960).

Process Hypotheses

7. If particular subjects dominate team deliberations and expend more time than their share, team decisions and performance may suffer despite use of the Perceptronics aid (Stein, 1975).

8. If the subject playing the role of group leader uses an autocratic leadership style, thus inhibiting team members from participating in deliberations and offering alternatives, team decisions and performance may suffer despite use of the Perceptronics aid (CACI, 1976).

9. If the experimenter's director interferes excessively in group processes or interprets data from the system or team preferences incorrectly, team decisions and performance may suffer despite use of the Perceptronics aid (Haines, et al., 1961).

Pre-Experimental Briefings

When the subjects arrive at the test site they receive two different briefings. The first is a training briefing, which consists of decision analysis instructions and hands-on experience with the decision aid for the experimental group, and problem-solving and team procedure discussions for the control group. The aided group should be instructed on the sensitivity procedures structured into the group aid at this point. For both group types, emphasis should be placed on the fact that the subjects will be working as a team in the upcoming experiments. At this stage, discussion among group members should be encouraged so that inhibitions are reduced and informal team procedures and rules of operation can be established.

The second briefing deals with the substance of the crisis management scenario and the team's tasking. Ideally, group members should read the written materials independently of each other and then be given an audio-visual briefing to summarize the salient facts. The written material for such viewgraphs was provided by CACI (1977). Discussion among group members concerning scenario facts should be prohibited until the decision-making session begins. Informational questions can be asked by team members, but responses should be limited to the data provided in the briefing materials to prevent bias in the distribution of information across teams.

Session Procedures

Once the session has begun, the experimenter should refrain from leading the team toward certain alternatives or joining in the discussion. Time checks should be announced every hour for each of the 3 hours, and a clock should be present so that the teams can pace themselves.

To assist in subsequent analysis of the group sessions (both aided and unaided), the experimenter should record a sequential list of all alternatives generated by the group, attributing each alternative to the initiating member and categorizing each alternative as either an action or event. This will assist in maintaining an audit trail of control group performance, as well as enable richer interpretation of aided group nodes in subsequent applications of the evaluation measures. A second observer in the control sessions might attempt to construct a decision tree of the generated alternatives as the discussion progresses. These data will save the experimenter from having to make several passes through the audio recordings of the sessions. Overall, special attention should be given to monitoring the unaided groups and making sure that there is an adequate audit trail of objectives discussed, alternatives generated and evaluated, and votes taken. Periodic photographs of blackboards and sequential numbering of paper used by control groups should also be attended to.

Post-Experimental Debriefing

Groups should be reminded that they must develop a final recommendation at the end of their 3-hour session. When the session is complete, the post-experimental attitude survey should be filled out by each member independently. Any verbal debriefing of the experiment by the experimenter should keep in mind whether the subjects will be undergoing a subsequent decision-making session (more on this in the Recommendations section below).

PROBLEMS AND RECOMMENDATIONS

The evaluation exercises that were conducted made certain problems with the group aiding system very apparent and the need for practical solutions critical. Several recommendations to reduce the impact of these problems are suggested below.

Timing

Timing is critical in crisis management decision-making situations. The team using the group aid in the current set of exercises was extremely passive in relation to the system, waiting for their alternatives or attributes to be entered by the intermediator and for prompts from the system to assign utilities or probabilities. Little discussion was carried on during these waiting periods despite prompts from the intermediator. Moreover, much time was expended by the intermediator during the first hour of the session in explaining each of the system's prompts (for attributes, alternatives, utilities, probabilities, and conflict resolution procedures), despite the fact that all of these elements had been described during the prior training session.

It is suggested that the training session for the aided team be more extensive, to the point where the team appears to be able to use the system entirely on its own. During the recent exercise, 20 minutes were spent on decision analysis instruction and one-half hour on group aid instruction. Perhaps another one-half hour is required for more hands-on experience with the aid. This should reduce the need for intermediator explanations during the session and reduce team passivity since members will be able to anticipate what the group aid requires them to do next. Of course, extending the instruction time for the aided group must be accompanied by increasing preparation time for the unaided group. Preparation time for both types of groups should be equivalent.

<u>Group Aid Limitations and Recommended Improvements</u>

Several structured elements of the group aid appeared to inhibit team performance. These can be rectified as suggested below.

1. When utility conflicts arise on alternatives, teams should be automatically shown the degree of differences before they are asked to decide on the procedure to resolve the conflict. These differences can be presented graphically.

2. As one of several conflict resolution procedures (in addition to MAUM and averaging), teams should be allowed to simply revote on the conflicting alternative subsequent to group discussion. This option would save much time over the MAUM procedure.

3. Groups should be enabled to override the sensitivity algorithm to access and expand action and event nodes that have been assigned high value by the group.

4. In order to accomplish this in an unbiased fashion, teams should automatically be shown the list of all nodes included in the tree with rolled-back values after each new node is added to the tree.

5. The hard copy audit trail produced by the system should have the computer clock print out the "time expended to generate alternatives at each node," as well as the "time expended assigning values and probabilities to those alternatives." This will aid in monitoring the percentage of time spent generating options and evaluating those options, respectively.

Experimental Design

The evaluation exercise indicated that unless very strict controls are taken in selecting subjects and minimizing intergroup differences, it may be difficult to validly compare aided with unaided group scores on the evaluation measures. In fact, it is probably impossible to control completely for team composition because not only must individual subjects be selected on the basis of specific criteria, but their assignment to teams results in the interaction of personalities and styles that cannot be completely anticipated or controlled.

Therefore, rather than the "posttest-only" control group design used in the current evaluation exercise, which is diagrammed below,[1]

$$\begin{array}{lll} R & X & O_1 \\ R & & O_2 \end{array}$$

[1] These designs are based on Campbell and Stanley (1963). In this and the following diagram, X = exposure to an experimental treatment, O = observation or measurement, and R = random assignment to conditions.

where one group experiences the aid and the other does not, a "pretest-posttest" control group design is recommended. It can be diagrammed as shown below,

R	0_1	X	0_2
R	0_3		0_4

Two groups are employed: One receives the aid treatment after undergoing an unaided session, and the second group undertakes two unaided sessions. This design controls for the following threats to validity: history, maturation, regression, and selection. Subjects must be randomized across experimental and control conditions, but matching may be used prior to randomizing. To control for subject mortality, the first and second decision-making sessions should be conducted on the same day. To control for learning and testing effects, the problem-solving scenarios given to both the experimental and control groups should be balanced. To implement this improved design, two scenarios must be developed of similar complexity and tasking.

Group Aid Utilization

Finally, it is recommended that the purpose for which the group aid was developed be reevaluated. As presently formatted, the aid provides little assistance to groups in generating alternatives and, in fact, may inhibit this process by implicitly limiting the number of alternatives that can be handled efficiently and intelligently in decision tree format by teams under pressure. On the other hand, the decision analysis algorithms in the group aid do appear to aid in evaluating and ranking the alternatives that are generated.

Optimal utilization of the group aid, it would seem, should take advantage of the structure it imposes on the alternative evaluation process, while ridding teams of its inhibiting effects during the alternative generation process. It is recommended that groups initially use proven idea generation and creativity processes such as brainstorming to develop a rich set of options. With these options in hand, the group can then employ the group aid to structure, rank order, and evaluate their alternatives by assigning values and probabilities. The separation of idea generation and evaluation procedures is likely to yield a rich and extensive set of alternatives as noted by Stein (1975). Overall, implementation of this recommendation will maximize the best elements of the group aid and may also yield higher quality decisions from participating groups.

BIBLIOGRAPHY

CACI (1978a) Further Development of Evaluation Methodologies for an Interactive Computer-Based Group Decision-Making Aid. Arlington, VA: CACI, Inc.-Federal.

_____ (1978b) Development of Evaluation Methodologies for the Perceptronics Interactive Computer-Aiding System for Decision-Making. Arlington, VA: CACI, Inc.-Federal.

_____ (1977) Simulated Briefing Materials: Operation TPCANYON. Arlington, VA: CACI, Inc.-Federal.

_____ (1976) The Impact of Computer-Based Decision Aids on Organizational Structure in the Task Force Staff. Arlington, VA: CACI, Inc.-Federal.

CAMPBELL, D.T. and J.C. STANLEY (1963) Experimental and Quasi-Experimental Designs for Research. Chicago: Rand-McNally.

GOLDMAN, M. (1965) "A Comparison of Individual and Group Performance for Varying Combinations of Initial Ability," Journal of Personality and Social Psychology 1, 3.

GREER, F., E. GALANTER, and P. NORDLIE (1954) "Interpersonal Knowledge and Individual and Group Effectiveness," Journal of Abnormal and Social Psychology 49.

HAINES, G., F. HEIDER, and D. REMINGTON (1961) "The Computer as a Small Group Member," Administrative Science Quarterly 6, 3.

LORGE, I. and H. SOLOMON (1960) "Group and Individual Performance in Problem Solving Related to Previous Exposure to Problem, Level of Aspiration, and Group Size," Behavioral Science 5, 1.

Perceptronics (1978) An Interactive Computer Aiding System for Group Decision Making. Woodland Hills, CA: Perceptronics Inc., Technical Report PQTR-1046-78-2.

STEIN, M. (1975) Stimulating Creativity: Volume II, Group Procedures. New York: Academic Press.