1·0
2·8   2·5
5·0   3·15 2·2
5·6
1·1   3·5   2·0
4·0
4·5   1·8
1·25 1·4   1·6

NATIONAL BUREAU OF STANDARDS
MICROCOPY RESOLUTION TEST CHART

TAEG

TAEG REPORT NO 32

DESIGN OF TRAINING SYSTEMS
THE DEVELOPMENT OF SCALING PROCEDURES

(12)
B.S

FOCUS ON THE TRAINED MAN

DECEMBER 1976

Technical Report: TAEG REPORT NO. 32

DESIGN OF TRAINING SYSTEMS

THE DEVELOPMENT OF SCALING PROCEDURES

ABSTRACT

This single volume final report expands the knowledge and procedural base
for making decisions where the potential outcomes of alternative decision
actions are characterized by multiple dimensions of value. This extends
and heavily relies upon an understanding of the work previously reported
in the Phase II-A Report.

The report includes extensive references to the current literature on
multi-attribute utility estimating in discussing both the theoretic as
well as the practical considerations faced by assessors and decision makers
in structuring data for decision-making purposes. Finally, a detailed set
of procedures is presented in Section II; it is intended to provide design
guidance in developing an interactive computer-based program of the Educa-
tional Technology Assessment Model (ETAM).

REPORT DOCUMENTATION PAGE		READ INSTRUCTIONS BEFORE COMPLETING FORM
1. REPORT NUMBER TAEG Report No. 32	2. GOVT ACCESSION NO.	3. RECIPIENT'S CATALOG NUMBER
4. TITLE (and Subtitle) Design of Training Systems: The Development of Scaling Techniques. *Procedures.*		5. TYPE OF REPORT & PERIOD COVERED Final 9/23 - 12/23/76 Rept. 23 Sep - 23 Dec 76
		6. PERFORMING ORG. REPORT NUMBER
7. AUTHOR(s) Robert B. Miller Larry R. Duffy		8. CONTRACT OR GRANT NUMBER(s) N61339-73-C-0097
9. PERFORMING ORGANIZATION NAME AND ADDRESS International Business Machines 7900 N. Astronaut Boulevard Cape Canaveral, Florida 32920		10. PROGRAM ELEMENT, PROJECT, TASK AREA & WORK UNIT NUMBERS
11. CONTROLLING OFFICE NAME AND ADDRESS Naval Training Equipment Center Training Analyis and Evaluation Group Orlando, Florida		12. REPORT DATE December 1976
		13. NUMBER OF PAGES 86 p.
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 is Unlimited.

DDC
RECEIVED
MAR 23 1978
B

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

18. SUPPLEMENTARY NOTES

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

Training Management Systems   Decision Theory
Computer-Based Models   Utility Theory
Educational Technology   Subjective Estimation
Training Technology

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

This single volume final report expands the knowledge and procedural base for making decisions where the potential outcomes of alternative decision actions are characterized by multiple dimensions of value. This extends and heavily relies upon an understanding of the work previously reported in the Phase II-A Report. (See AD-A015 295)

DD FORM 1 JAN 73 1473   EDITION OF 1 NOV 65 IS OBSOLETE
S N 0102-LF-014-6601

EB

408 510

20. ABSTRACT (Cont'd)

The report includes extensive references to the current literature on multi-attribute utility estimating in discussing both the theoretic as well as the practical considerations faced by assessors and decision makers in structuring data for decision-making purposes. ~~Finally~~ a detailed set of procedures is presented in Section II; it is intended to provide design guidance in developing an interactive computer-based program of the Educational Technology Assessment Model (ETAM).

ACCESSION for

NTIS	White Section
DDC	Buff Section ☐
UNANNOUNCED	☐
JUSTIFICATION	

BY
DISTRIBUTION/AVAILABILITY CODES

Dist.	AVAIL. and/or SPECIAL
A	

DESIGN OF TRAINING SYSTEMS

THE DEVELOPMENT OF SCALING PROCEDURES

Robert B. Miller

Larry R. Duffy

This Study Was Performed By

International Business Machines Corporation

for the

Training Analysis and Evaluation Group

December 1976

ALFRED F. SMODE, Ph.D., Director
Training Analysis & Evaluation Group

FOREWORD

This report presents the final results in the development of scaling procedures to be employed as a part of the overall ETAM Procedures contained in the ETAM I, Phase II-A Final Report. This report represents the output for Task 1 of Phase II-B, which is a part of a multiphase effort called "Design of Training Systems," undertaken in consonance with the requirements of Advanced Development Objective 43-03X, "Education and Training."

Sincere thanks is expressed for the cooperation of the Director, Dr. A. F. Smode, and members of the U. S. Navy Training Analysis and Evaluation Group for making available certain documentation and for providing assistance in various areas of the project. Acknowledgment is also made to Bruce R. Judd of the Decision Analysis Group at the Stanford Research Institute, to William Giauque of the Naval Postgraduate School, and to Detlof von Winterfeldt of the Social Science Research Group at the University of Southern California, who supplied some invaluable papers and references as well as provided personal communication on assumptions and simplifications.

The scaling procedures and associated rationale were prepared by Dr. R. B. Miller and Mr. L. R. Duffy. Mr. R. E. Hallman provided management for the project; Mrs. C. Reilly provided secretarial and art services.

TAEG REPORT NO. 32

TABLE OF CONTENTS

SECTION I

TASK INTRODUCTION

BACKGROUND ON ETAM I

This report is a technical extension to the Educational Technology Assessment Model (ETAM I) and presents procedures, formats and rationales for converting a mass of benefit analysis data and cost analysis data into decision-making presentations. The promised benefits and costs are expressed in a number of variables and magnitudes that are not directly commensurate with each other. The scaling procedures facilitate the assessor, expert in translating these several kinds of benefits, liabilities and costs, into a single continuum of worth and arrive at a single figure of merit for each projected outcome of decision alternatives for accepting or rejecting the proposed innovation. The scaling procedures separate and then integrate subjective probability estimates for success and failure outcomes, given the selection of an alternative. The scaling procedures also enable the assessor to express judgments of relative importance of one benefit (or liability) as compared with others.

Plans call for the use of a "conversational" computer terminal to implement the scaling and conversion operations, as well as for the antecedent operations that develop assessment data. The general ETAM philosophy is to offer the human user computer facilities for automatic functions in data reduction, conversion and calculation, but not to force these functions if the user prefers to use his own intuitive or judgmental capabilities. Generally, the human is required to make common sense audits of all steps in data conversions even when they are semiautomatic. In short, the computer remains a tool for the user; it is never intended as a surrogate for human judgment in generating assessment data, reducing it to a decision map, or in making the decision itself.

The present report is more understandable if the reader has a picture of ETAM I, described in another publication in this series. The following description is a conceptual rather than a procedural outline of ETAM I.

The Educational Technology Assessment Model (ETAM I) is a set of conceptual and procedural structures for the purpose of making technical and business decisions about proposed innovations or changes to any aspect of training. The proposed change may range from the manner of deriving and specifying training objectives, through methods of instructional delivery and student training environments. ETAM attempts to maximize the effective use of all information available about the change or innovation at any given stage of inquiry. This includes the ability of experts to make interpretations, and estimate probabilities of successfully introducing the innovation. Innovation in the ETAM context refers to any change which can be justified, based upon reasonably acceptable criteria, and planned for within the existing planning framework (e.g., Programming, Planning and Budgeting System).

ETAM consists of a set of procedural formats for describing the proposed innovation, determining its potential range of applicability in the entire training establishment and job-task repertory; for structuring the pattern of comprehensive unit costs relevant to the innovation; and for organizing and presenting the benefit and cost data. The presentation is both in profiles of source data, and in the form of decision trees that can facilitate rational decision making in the adoption or rejection of the innovation, or in deferring the decision pending specific study outcomes. All aspects of the procedure and the data are open to the decision maker who can thereby use intuitive processes to challenge, deny, or confirm the presentation of results by the business analysis model.

The use of ETAM procedures requires an expert of the consultant type to relate the formalized description of the innovation -- and any data available about it -- to the full range of instructional devices, courses and course objectives, jobs and job-tasks, and student characteristics. The classification system that guides the assessor is a comprehensive, procedurally-oriented taxonomic structure of processes and products in the economics of instruction and in the psychology of learning. The classification scheme enables the assessor to select descriptors that apply to the description of the innovation. These descriptors may be used to search data bases (in a computer, or on paper or in people's heads) for relevant job-tasks, courses and course objectives, instructional vehicles, and student characteristics that are specifically applicable to the innovation. The ETAM reference material is a comprehensive, conceptual "model" of the technology of training and learning. The applicable "range-of-effect" of the innovation generates the parameters and multipliers (such as number of students, courses and training hours) for quantifying the benefits and costs attributable to the innovation.

In the first stage of assessment, a preliminary technical feasibility analysis is performed. If the innovation survives this stage, its full potential range of effect in benefits and costs is assessed. Magnitudes of benefits -- and liabilities -- are estimated, as are subjective probabilities of achieving the effects. This determination of range, magnitude and probabilities of effects is the second major phase of assessment. The third phase is computational. It consists of a business analysis of costable benefits and liabilities, and considers opportunity costs in the generation of a return on investment analysis adjusted for risk. The model yields decision trees that enable the executive decision maker to perceive comparative risks and returns on investment from (a) adopting the innovation now; (b) deferring the decision until specific further information is obtained or studies completed; or (c) rejecting the innovation. The decision maker may set aside working with the decision-tree structure if he prefers to work with the profiles of source assessment data, which are presented to him as supplement, or alternative, to the decision tree.

Conceptually, the ETAM process can be viewed as diverging to embrace identification of the full range of application, then converging on a decision to accept or reject the innovation considering a more precisely defined use target. Figure I-1 portrays this conceptual view of ETAM. A more detailed view of the entire ETAM procedural sequence is shown in Figure I-2.

ETAM cannot generate data that do not exist. But it does make the best and most exhaustive use of available information about the innovation, of expertise, and of state-of-the-art in instructional technology. But source data

FIGURE I—1. ETAM CONCEPTUAL OVERVIEW

TASK 4
PERFORM ANALYTIC FEASIBILITY ASSESSMENT

TASK 3
MAKE PRELIMINARY FEASIBILITY PROFILE

TASK 6
PERFORM COST-BENEFITS ANALYSIS

TASK 5
DETERMINE RANGE-OF-EFFECT

TO IMPLEMENTATION PLANNING PHASE

TO CONTINUED STUDY PLANNING PHASE

END

EDUCATIONAL TECHNOLOGY ASSESSMENT MODEL (ETAM)

PROPOSED POTENTIALLY TRAINING-RELATED INNOVATION

INNOVATOR LITERATURE SEARCHES EXPERTS IN FIELD

TASK 1
FORMALIZE THE DESCRIPTION OF THE INNOVATION

TASK 2
DEVELOP/EXAMINE ALTERNATIVES TO THE INNOVATION

TASK 8
MAKE ACCEPT/REJECT/STUDY DECISION
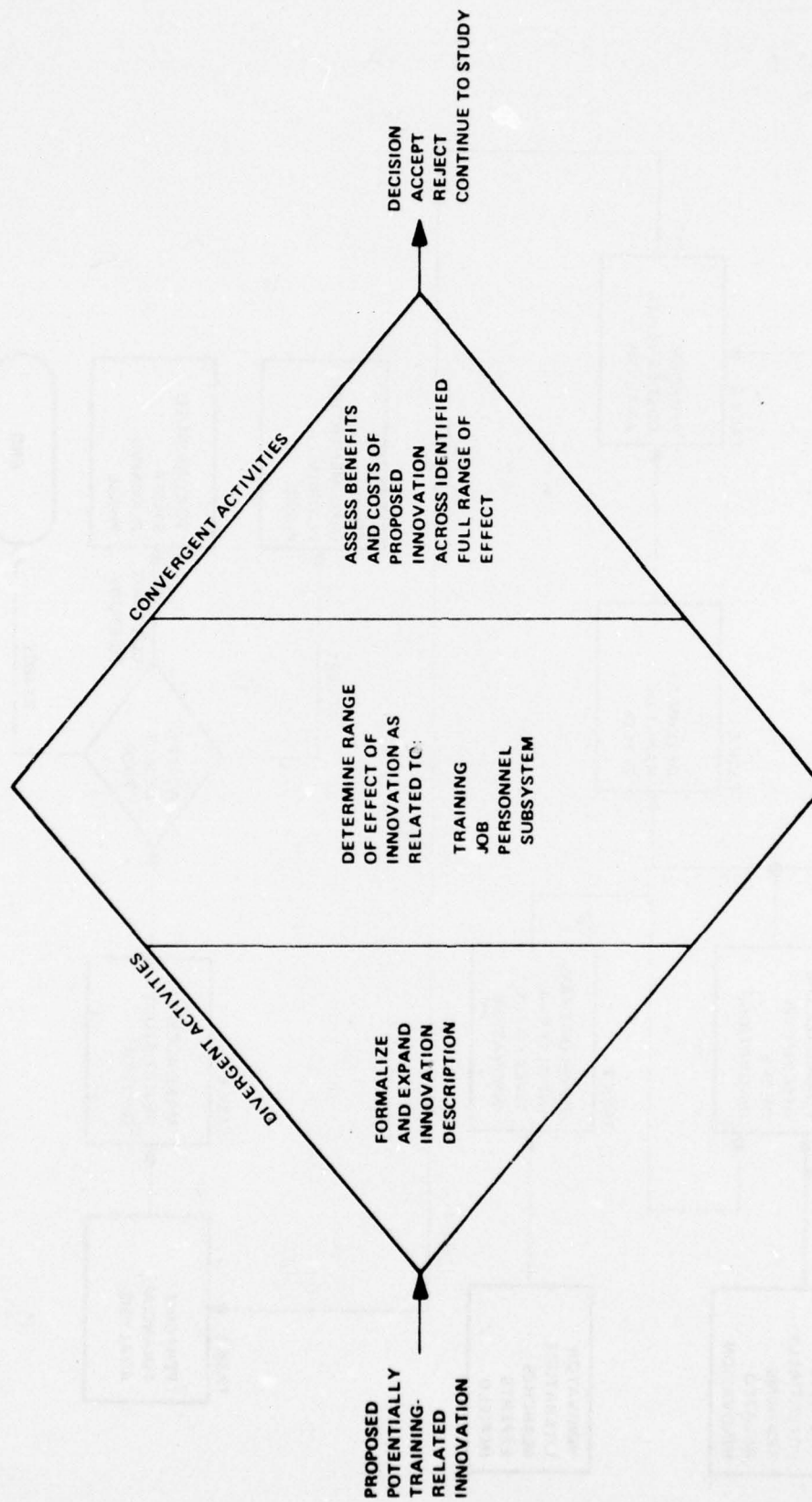
TASK 7
PERFORM FINANCIAL ANALYSIS

ACCEPT/REJECT PROP

ACCEPT
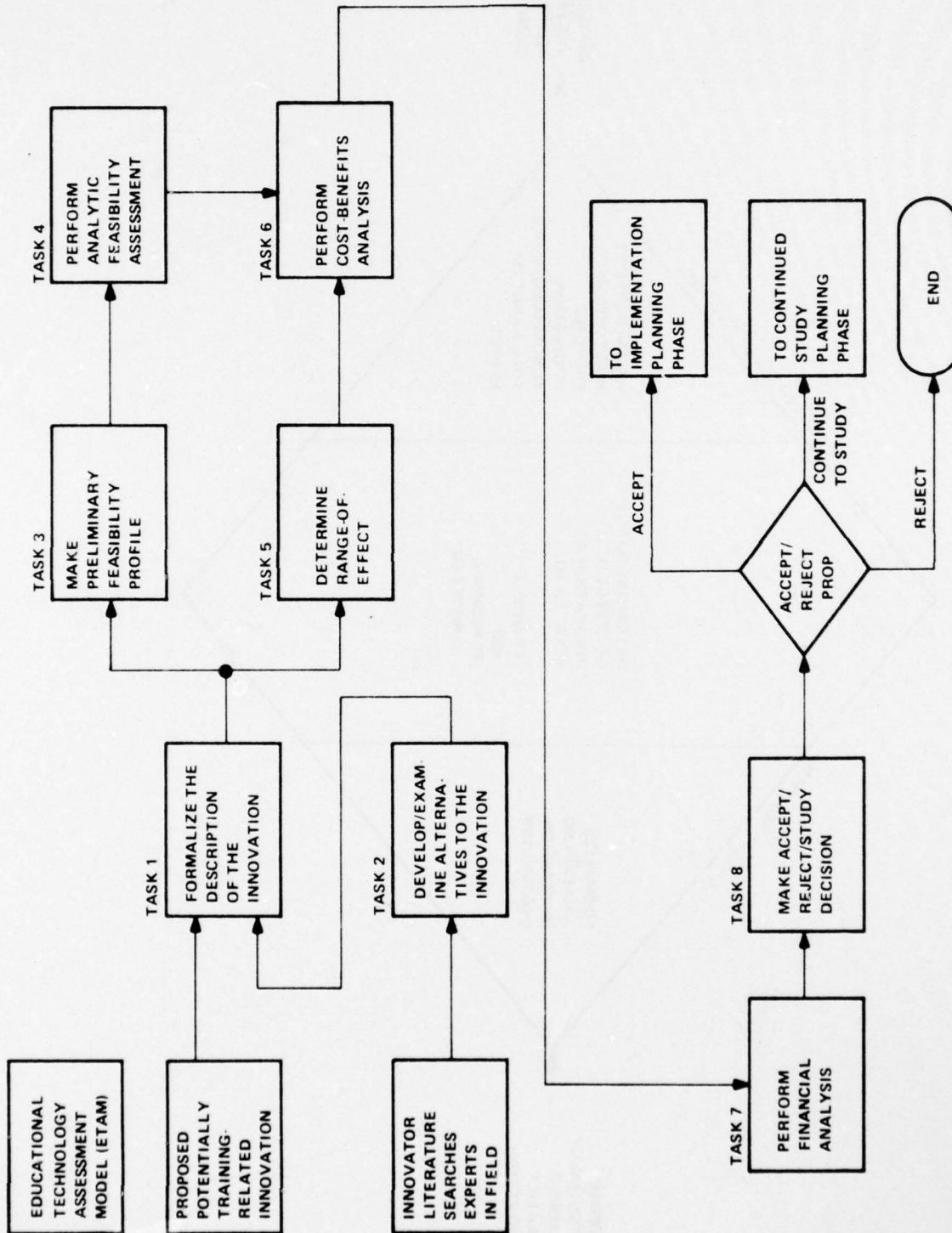
CONTINUE TO STUDY

REJECT

FIGURE I–2. ETAM PROCEDURAL SEQUENCE

are measurements where available and are guided human judgment where data are not available or must be extrapolated. The decision maker is presented with summarized pictures of outcomes in benefits, costs and expressions of the assessor's uncertainties about his projections. The model also highlights sensitive variables that justify focused inquiry in borderline decisions. It also suggests the potential worth of a given R&D expenditure.

The consultant-level specialist in Navy training should be able to learn to apply the ETAM procedures to innovations, even difficult ones to classify in terms of range of effect, within a couple of weeks of intensive study.

The first phase of the ETAM project confirms that this form of technology assessment is feasible as a procedure up to and including the level of sophisticated decision making. Implementation of the models for training and application of the procedures with interactive computer terminals require several targetted follow-on developmental studies. The analytic techniques for the training enterprise may be generalized to the appraisal of quanti- tative benefits in dollars and other values to be obtained by increases of skill and motivation in on-the-job performance.

The ETAM effort focused on two key areas considered essential to the success- ful operation of a final model product. These were:

1) Taxonomic structures and search techniques for generalizing an innovation in terms of its potential full "range-of-effect."

2) Assessment procedures and decision-making constructs for training management to use in reaching valid conclusions about an educa- tional innovation based on its overall cost and benefit charac- teristics.

The reader is recommended to see the ETAM I report itself for content under these two topics.

STATEMENT OF SCALING PROBLEM

ETAM I developed scales that enable the assessor to express subjective estimates of relative "Importance" of an innovation and its effects. In ETAM I a rough scale of "Importance" was provided for the assessor. The level of importance could be interpreted subjectively by the decision maker from this scale. The ETAM I specification for a quantititive model includes as a processing input, the variable "Degree of Importance" as a multiplier for benefits and costs as a weighting factor. While the factor was included as a multiplier within the formal model structure, the con- version between the verbal expression of importance to valid and reliable model parameters has not been formulated. Because estimates of "importance" must be anchored in contexts, it is essential that the procedure facilitate the user's identification and specification of the context relevant to the decision of relative importance and shared by the decision maker.

A further aspect of the scaling problem relates to the appropriate estima- tion of the utility of some outcome reached as a result of a decision alternative, especially where the outcome's utility is characterized by multiple attributes, each having an associated level of utility. Estima- tion methods must handle variables (also known as attributes) which are

both qualitative as well as quantitative in nature. Also, models must be defined for combining into a single figure of merit (utility) the individual utilities estimated at each outcome for these types of variables.

A similar estimate scaling problem applies, perhaps to a lesser degree, to the assessor's expression of "probability of success" of the innovation which is also a multiplier (at a different level) of values expressed as projected costs and benefits. The ETAM decision tree shown in Figure I-3 describes the context for utility and probability estimations which gave rise to the scaling task. Outcomes A, B, C and D are characterized by a multi-attributed utility pattern while junctures U, V, W, etc. require estimates of probabilities.

APPROACH TO THE SCALING TASK

The purpose for the scaling task was to develop the necessary estimation techniques for incorporation into the previously developed ETAM procedural model. These techniques are intended to increase the reliability and validity of subjective estimates of relative importance and of outcome utility. A rather extensive body of theory exists in the areas of subjective estimating and multi-attributed utilities. The literature also contains substantial references to experiments applying the theory to practical real-world situations. This task, therefore, involved significant effort in reviewing and assimilating the literature in order to derive practical methods for estimating utility which could be incorporated into ETAM. The rationales for the procedures are embodied in this report, and several scenarios employing facets of the procedures are included to demonstrate their practical use.

An underlying aspect of this task was to develop sufficient guidance in the procedural outline and accompanying scenarios to serve as a design specification for programming this facet of ETAM for interactive use.

SUMMARY

The task objectives involved the development of scaling techniques, the preparation of supportive and exemplary scenarios, and the documentation of selection and rejection rationale. The entire process is targetted at improving the overall validity of the ETAM procedures by:

1) Improving the selection of a comprehensive set of relevant decision variables (attributes).

2) Giving additional insight to the assessor in evaluating benefit patterns characterized by a number of attributes of differing levels of importance.

3) Providing means for extracting maximum information content from the available data.

4) Facilitating the use of multiple participants in the assessment process -- to bring to bear on the problem diverse areas of expertise in a coordinated fashion.
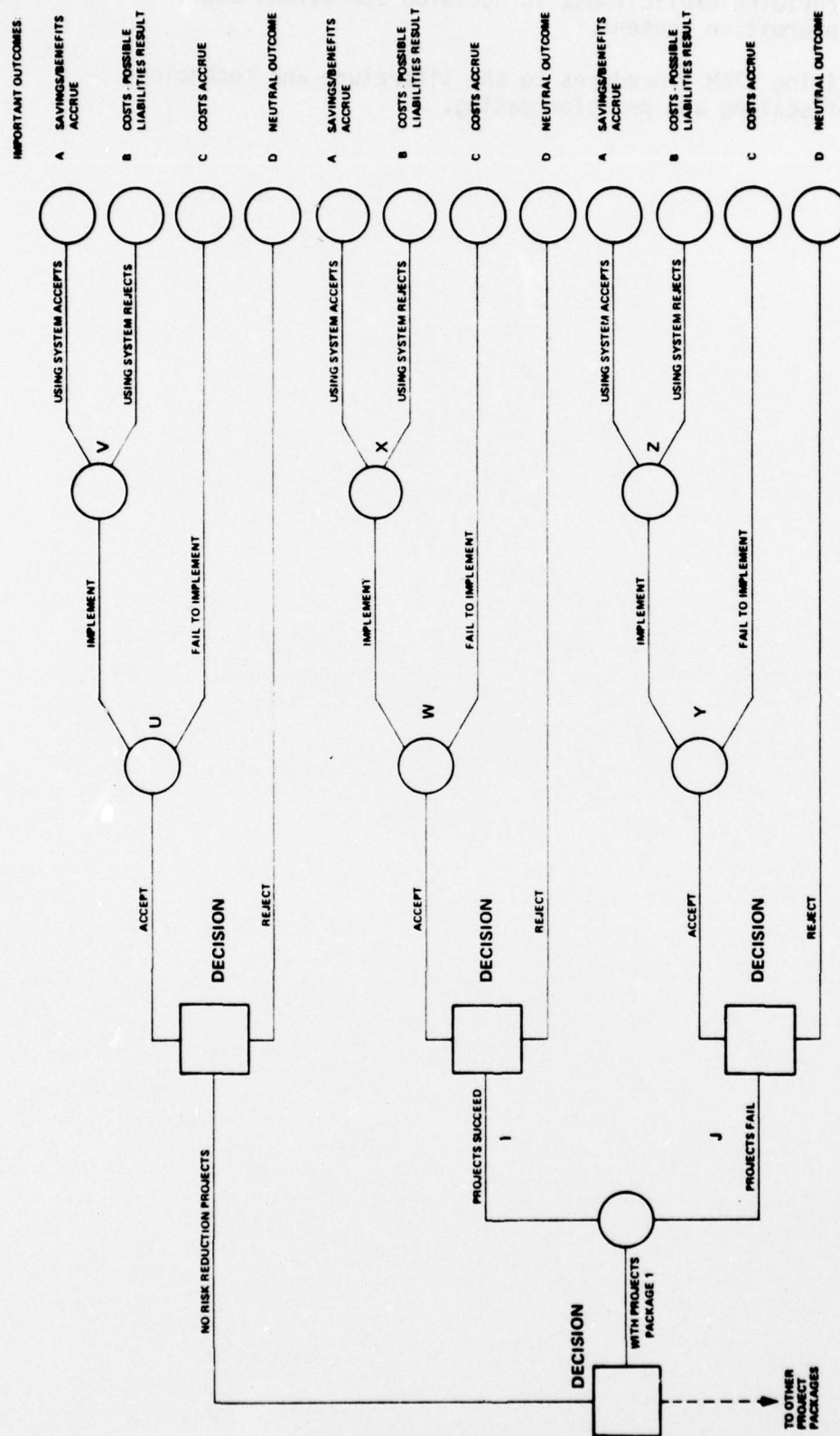
FIGURE I-3. DECISION TREE FOR ASSESSMENT OF TRAINING INNOVATION

5) Providing explicitness in decision operations and
    information content.

6) Tieing ETAM procedures to the literature and technology
    of scaling and decision making.

SECTION II

SCALING PROCEDURES FOR THE RANGE-OF-EFFECT/ASSESSMENT INTERFACE

INTRODUCTION

Any innovation being considered for assessment will generally provide a multiplicity of benefits (or liabilities - here considered a negative benefit) if successfully implemented. Each benefit is characterized by name and by magnitude. The name of a benefit also may be called a relevant variable or attribute. Its magnitude may initially be stated in one or more of several ways. If the benefit is quantifiable and has some fairly well defined unit of measurement, then magnitude may be expressed as a percentage change from some baseline value (e.g., a ten percent increase in attrition), or as some absolute change in number of units (e.g., 400 more attritions per year). Qualitative benefits may require magnitude to be stated in a less concrete way (e.g., a considerable improvement in student morale). However a benefit's value might be visualized, it is the purpose of these procedures, along with the exemplary scenarios, to provide a method for the assessor of an innovation to quantify the utility associated with a group of benefits considering various outcomes of several alternative decision paths. The following definitions are important to the understanding of these procedures, and of the rationale that support them.

DEFINITIONS

1. Alternative

    The action path(s) that results from a specific decision. An alternative may lead to one of several outcomes, each with an estimated probability of occurring.

2. Outcome

    The projected end result of a decision. An outcome may be projected as a result of 1) a decision, and 2) a series of probabilistic occurrences. If an innovation were accepted (by decision), one outcome may be that it was successfully implemented (which originally had some estimated probability of happening) and that the using system was improved as a result of the innovation being successfully implemented.

3. Benefit

    A relevant variable, impacted by the decision about the innovation, which can assume various magnitudes or states for each of the possible outcomes which may result from alternative decision paths.

4. Attribute

    Another term for the relevant variables to be evaluated for each of the possible outcomes identified within a decision tree.

5. Innovation

Any change which can be planned for and evaluated on some acceptable set of criteria.

6. Utility

Utility describes a level of preference for something. The measure of utility will generally be in some arbitrary units; e.g., on a scale of 0 to 100. In the context of the ETAM Procedures, the expression of utility will be derived from a pattern of organizational goals, as opposed to being based upon some personal pattern of preferences. Utility units are, therefore, units of organizational or system preference.

7. Value

Value, when used in its strictest sense, refers to an objective measurement dimension such as dollars, weight, time, etc. It differs from utility in that points on a value scale may not linearly reflect the preference for those points. If one prefers a second of two cars only half as much as the first, and the value of each car is $5,000 with an arbitrary measure of utility for the first car being 50, then the total value of the two cars will be $5,000 + $5,000 = $10,000 while their combined utility will only be 50 + .5 x 50 = 75 (rather than 100).

8. Expected Value

This is the sum of all values factored by their probability of being acquired. If there is a 70% chance of acquiring $1,000 and a 30% chance that only $100 will be obtained, the expected value will be $730. (.70 x $1,000 + .30 X $100).

9. Expected Utility

Mathematically, expected utility is calculated the same as expected value, except that units of utility have been substituted for the value figures.

10. Worth

Worth has been used in these procedures as a synonym for utility. It states a level of preference for something.

NOTATIONAL CONVENTION

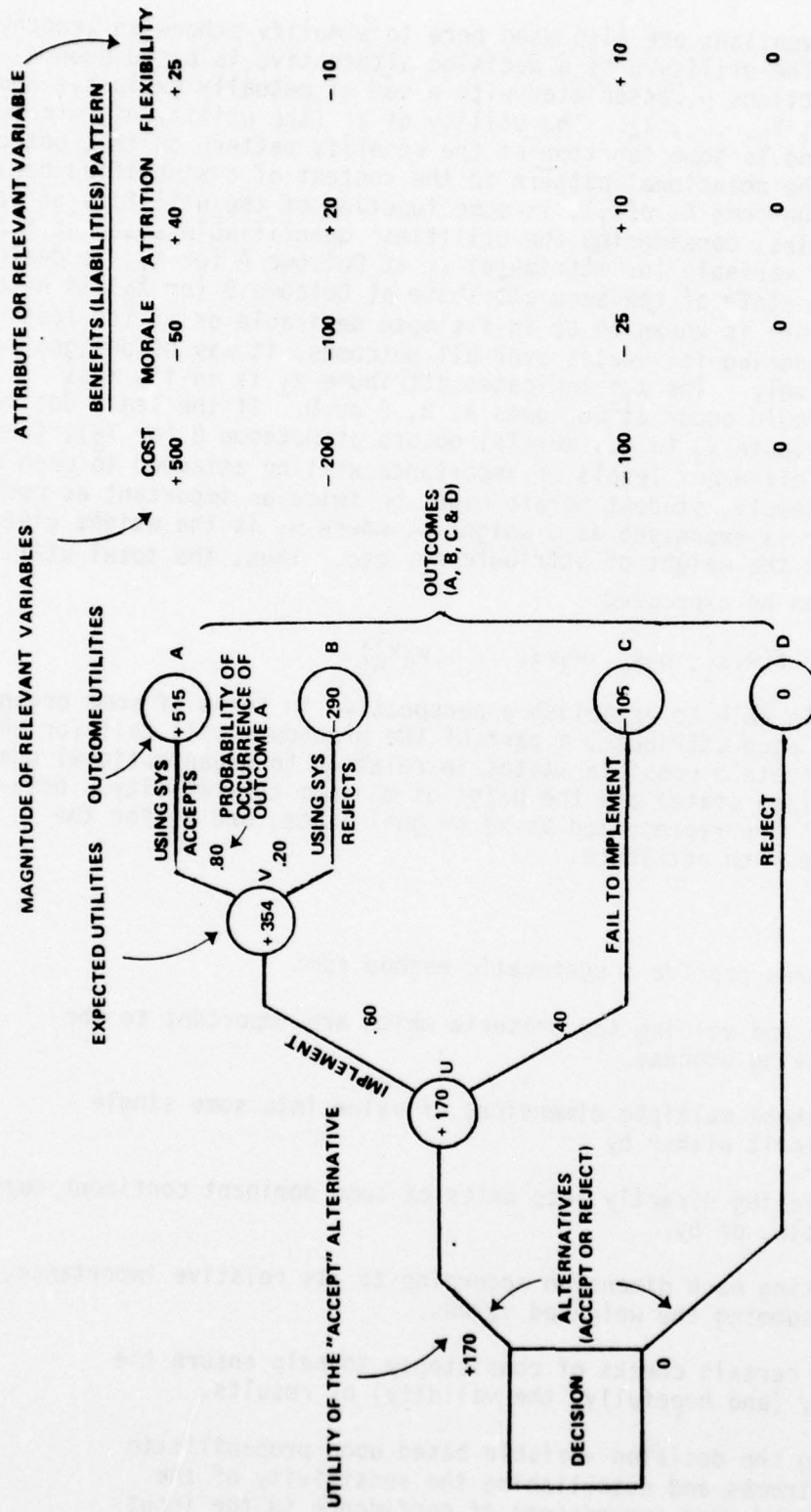The decision tree in Figure II-1 incorporates illustrations of many of the previously defined terms.

FIGURE II–1. DECISION TREE TERMINOLOGY

Certain notational conventions are also used here to simplify otherwise lengthy verbal descriptions. The utility U of a decision alternative is based upon a set of probability functions p, associated with a set of mutually exclusive and exhaustive outcomes X_1, X_2, ..., X_n. The utility of X_1 (the utility of Outcome A), is defined $U(X_1)$ and is some function of the benefits pattern on that outcome. Figure II-2 provides the notational pattern in the context of a simplified decision tree. The utility at Outcome A, $U(X_1)$, is some function of the utilities of each of the relevant variables, considering the utilities' quantifiable state at that outcome. The state of variable (or attribute) x_1 at Outcome A (or X_1) is designated x_1^1, likewise the state of the same attribute at Outcome B (or X_2) is noted as x_1^2. When an attribute is known to be in its mose desirable or at its least desirable state, considering its states over all outcomes, it may be designated as x_1^* or x_{1_*} respectively. The x_1^* indicates attribute x_1 is in its most desirable state, and could occur at Outcomes A, B, C or D. If the least desirable state of the attribute x_3 (e.g., morale) occurs at Outcome B (or X_2), then $x_{3_*} = x_3^2$. Generally, different levels of importance will be attached to each of the attributes; for example, student morale might be twice as important as cost. This importance factor is expressed as a weight w, where w_1 is the weight given to attribute x_1, w_2 is the weight of attribute x_2, etc. Thus, the total utility at some outcome X_i can be expressed:

$$U(X_i) = f(w_1 x_1^i, w_2 x_2^i, w_3 x_3^i, \ldots, w_n x_n^i)$$

So that the assessor is able to establish a perspective in terms of some organizational standard for each attribute, a part of the procedure will call for the estimation of an attribute's possible states in relation to organizational goals (a more or less idealized state) and the point of minimum tolerability. Notationally, these states are represented as x_i^g or goal-state, and x_i^t for the tolerable-state of the i-th attribute.

PROCEDURAL ABSTRACT

The following procedures provide a systematic method for:

a) Identifying and valuing the criteria which are important to the decision making process.

b) Combining these multiple dimensions of value into some single figure of merit either by

    1) Translating directly into units of some dominant continous base variable, or by

    2) Weighting each dimension according to its relative importance, then summing the weighted values.

c) Performing certain checks of consistency to help ensure the reliability (and hopefully, the validity) of results.

d) Determining the decision variable based upon probabilistic sets of outcomes and establishing the sensitivity of the variable based upon expressions of confidence in the input data.

$$U(X_1) = f(w_1 x_1^1, w_2 x_2^1, w_3 x_3^1, \ldots, w_n x_n^1)$$

$$U(X_2) = f(w_1 x_1^2, w_2 x_2^2, w_3 x_3^2, \ldots, w_n x_n^2)$$

$$U(X_3) = f(w_1 x_1^3, w_2 x_2^3, w_3 x_3^3, \ldots, w_n x_n^3)$$

$$U(X_4) = f(w_1 x_1^4, w_2 x_2^4, w_3 x_3^4, \ldots, w_n x_n^4)$$

A

USING SYS
ACCEPTS

B

USING SYS
REJECTS

C

FAIL TO IMPLEMENT
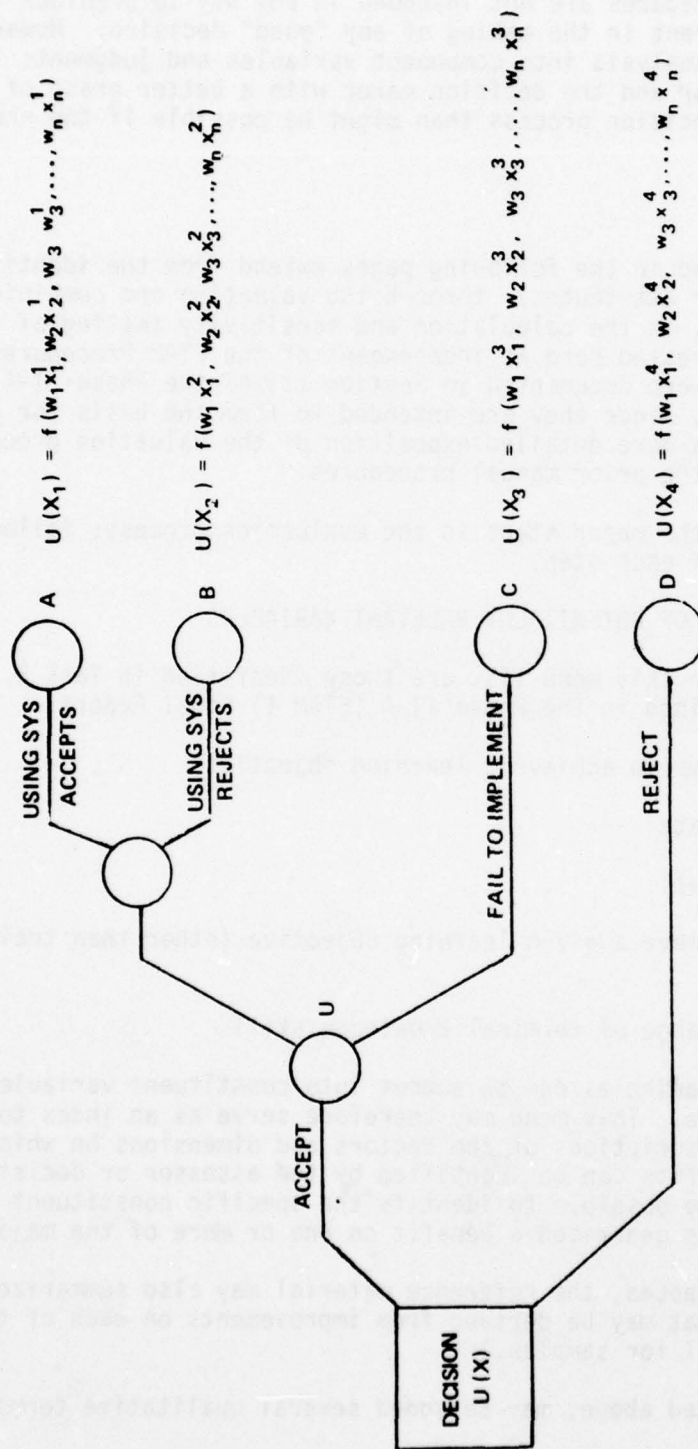
D

REJECT

U

ACCEPT

DECISION
U (X)

FIGURE II–2. DECISION TREE NOTATIONAL CONVENTIONS

Where appropriate, cautionary notes and references have been included. They are primarily intended to point out certain assumptions on which the procedures are based. These procedures are not intended in any way to preclude the judgmental processes inherent in the making of any "good" decision. However, the decomposition of the analysis into component variables and judgments is expected to provide the assessor and the decision maker with a better grasp of the relevant factors in the decision process than might be possible if the problem were treated wholistically.

PROCEDURAL FLOW

The procedures outlined in the following pages extend from the identification of relevant variables (or attributes), through the valuation and combining of attributes by outcome, to the calculation and sensitivity testing of the decision variable. They are treated here as independent of the ETAM Procedures previously developed, and which were documented in Section III of the Phase II-A (ETAM 1) Final Report; however, since they are intended to form the basis for programming ETAM, they represent a more detailed exposition of the valuation process and, therefore, supercede the prior manual procedures.

Figure II-3 presents the major steps in the evaluation process; following is a general description of each step.

Step 1 - PRESENT MENU OF POTENTIALLY RELEVANT VARIABLES

The major variables in this menu list are those identified in Task 5.18 of the ETAM Procedures contained in the Phase II-A (ETAM I) Final Report.

> Training time to achieve a learning objective
>
> Attrition rate
>
> Aptitude level
>
> Cost to achieve a given learning objective (other than training time)
>
> Level and range of terminal knowledge/skill

Each of these major variables can be subset into constituent variables, most of which are quantitative. This menu may therefore serve as an index to other menus, and to text descriptions of the factors and dimensions on which cost benefits and process benefits can be identified by the assessor or decision maker. It should therefore be possible to identify the specific constituent factors in an innovation that has generated a benefit on one or more of the major variables.

For instructional purposes, the reference material may also summarize the kinds of system goodness that may be derived from improvements on each of these variables. (See Appendix I for samples.)

To the small list cited above, may be added several qualitative terms, such as:

> Flexibility in operations
>
> Student motivation to learn

INPUTS

STANDARD LIST BY TRAINING, JOB, AND EQUIPMENT CATEGORIES

INNOVATION DESCRIPTION, OPERATIONAL CONTEXTS FOR ALL VARIABLES

PRESENT OR BASELINE STATE OF SYSTEM INCORPORATING IDENTIFIED ATTRIBUTES

ORGANIZATIONAL CONTEXT, I.E., GOALS. BASELINE STATE OF SYSTEM WITH RESPECT TO GOALS

RELATIONAL DESCRIPTION OF ATTRIBUTES AND GOAL PRIORITIES

TRADED-OFF OR WEIGHTED ATTRIBUTE VALUES FOR EACH OUTCOME

ATTRIBUTE AND OUTCOME UTILITIES

PROBABILITY ESTIMATES, RISK REDUCTION PROJECTS FROM ETAM TASKS 3, 4 AND 6

UTILITIES OF OUTCOMES, EVENT PROBABILITIES

EXPRESSION OF CONFIDENCE FOR EACH VARIABLE

INFORMATION AND RATIONALE ASSUMPTIONS

PROCESS STEP

1 — **PRESENT MENU OF RELEVANT VARIABLES (ATTRIBUTES)**

2 — **SELECT AND/OR IDENTIFY RELEVANT ATTRIBUTES**

3 — **CREATE OUTCOME SCENARIOS USING ALL ATTRIBUTES**

4 — **PREPARE DESCRIPTIVE (GRAPHICAL) REPRESENTATION OF ATTRIBUTES**

CONTINUOUS BASE ATTRIBUTE ? — YES / NO

5 — **CONVERT ATTRIBUTES INTO UNITS OF A CONTINUOUS BASE ATTRIBUTE**

7 — **ASSIGN IMPORTANCE LEVEL TO EACH ATTRIBUTE**

6 — **TRANSFORM BASE ATTRIBUTE VALUES INTO SYSTEM UTILITIES**

8 — **TRANSFORM WEIGHTED ATTRIBUTES INTO SYSTEM UTILITIES**

9 — **PERFORM CONSISTENCY CHECKS OF OUTCOMES AND OF PROCESS**

10 — **FINALIZE DECISION TREE AND COMPUTE DECISION VARIABLE**

11 — **PERFORM SENSITIVITY ANALYSIS**

12 — **PREPARE INPUT TO DECISION MAKER**

13 — **MAKE DECISION** — REJECT / ACCEPT

ACCEPT WITH PROJECTS

OUTPUTS

PRESENTATION OF ATTRIBUTES TO ASSESSOR

LIST OF RELEVANT ATTRIBUTES APPROPRIATE FOR EVALUATION

DEFINITION OF INCREMENTAL STATE CHANGE FOR EACH ATTRIBUTE AT EACH POTENTIAL OUTCOME

GRAPHICAL PRESENTATION OF EACH ATTRIBUTE'S LEVEL WITH RESPECT TO GOAL STATE.

VALUE OF EACH ATTRIBUTE LEVEL FOR EACH OUTCOME

TOTAL UTILITY OF EACH OUTCOME EXPRESSED IN COMMON UNITS

MEASURE OF RELIABILITY, NECESSARY CORRECTIONS

UTILITY FIGURE OF MERIT BY ALTERNATIVE

SENSITIVITY OF DECISION TO CHANGE. SUGGESTED NEED FOR ADDITIONAL INFORMATION

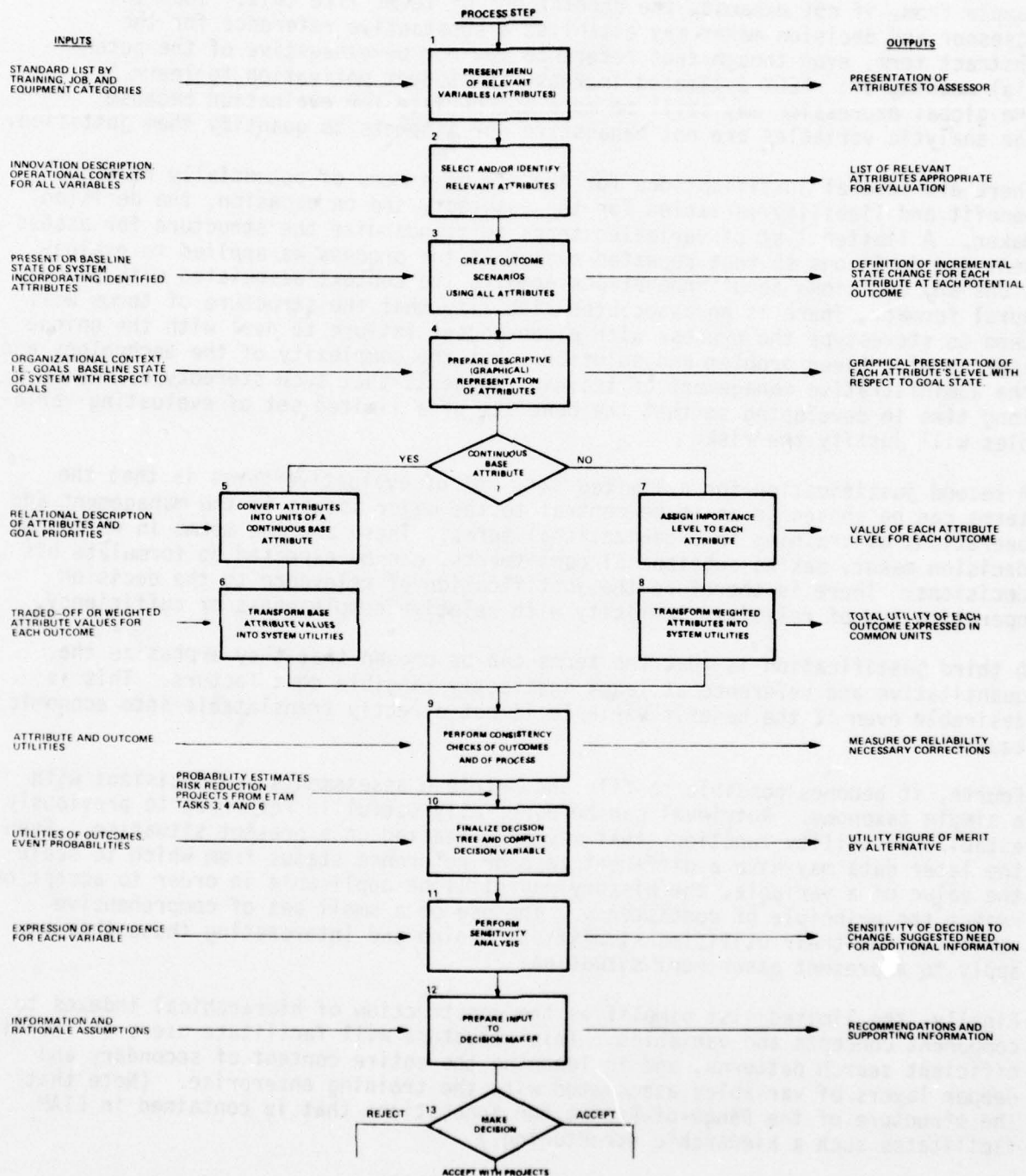**RECOMMENDATIONS AND SUPPORTING INFORMATION**

FIGURE II–3.   MAJOR STEPS IN EVALUATION PROCESS

The user should be able to use terms such as this as an index to lists and descriptions of operations, preferably quanitifiable, that will extensively sample from, if not exhaust, the denotations of terms like this. Thus the assessor and decision maker may establish a substantive reference for the abstract term, even though that reference may not be exhaustive of the potential meaning of: "20% estimated increase in student motivation to learn." The global expression may still be more appropriate for evaluation because the analytic variables are not exhaustive nor attempts to quantify them justified.

There are several justifications for the use of a menu of potentially relevant benefit and liability variables for the assessor, and on occasion, the decision maker. A limited list of variables tends to standardize the structure for assessment and decisions so that repeated cycles of the process as applied to evaluations and decisions about innovations acquire the context associated with procedural format. There is an associated liability that the structure of terms will tend to stereotype the process with a consequent failure to deal with the unique aspects of a given problem and solution. But the complexity of the technology and the administrative management of training suggests that such stereotypes may be a long time in developing so that the benefits of a limited set of evaluating variables will justify the risk.

A second justification for a limited menu set of evaluative terms is that the terms can be chosen so as to be central to the major issues in the management and operations of training in organizational terms. These are the terms in which the decision maker, making substantial commitments, can be expected to formulate his decisions. There is therefore the justification of relevance to the decision operation and of relative simplicity with relative completeness or sufficiency.

A third justification is that the terms can be chosen that they emphasize the quantitative and reference at least indirectly tangible cost factors. This is desirable even if the benefit variable is not directly translatable into economic equivalents.

Fourth, it becomes possible to file and retrieve assessments and decisions with a simple taxonomy. Retrieval can be especially useful in referring to previously established utility functions that may have bearing on a present situation. Even the later data may have a different zero or reference status from which to scale the value of a variable, the history may still be applicable in order to accept or reject the principle of consistency. The use of a small set of comprehensive variables and their utilities simplifies finding and interpreting those that may apply to a present assessment situation.

Finally, the limited list simplifies the construction of hierarchical indexes to component concepts and variables. This structure will facilitate users in learning efficient search patterns, and in learning the entire content of secondary and deeper layers of variables associated with the training enterprise. (Note that the structure of the Range-of-Effect for innovations that is contained in ETAM facilitates such a hierarchic structuring.)

Step 2 - SELECT AND/OR IDENTIFY RELEVANT VARIABLES

The assessor selects those variables from the menu list which fit the benefits and liabilities contained in the innovation. The variables selected, of course, will

be related to the types of entities affected by the innovation; i.e., courses, vehicles, or jobs.

Any list of cost or operating variables is based on what is expected to work with acceptable efficiency and effectiveness. It is heuristically derived. It should therefore be subject to being changed either by addition or deletion on the basis of need and experience. It is possible, for example, that the concept "career versatility" becomes important to training operations. Presumably the term can be given operational characterizations which, if not exhaustive, will cover its major meaning. The concept may become a new variable that should be added to the list of benefit/liability variables. A new term is added to the menu, and perhaps an old one is dropped as no longer useful.

As a general rule, a menu list of more than a dozen attributes presented as a unit from which to make conceptual (as contrasted with procedural) discriminations and selections becomes unwieldy for the user.

Step 3 - CREATE OUTCOME SCENARIOS USING ALL ATTRIBUTES

A resulting state for each relevant costable variable should be structured for each outcome to be considered. These states should be considered as deterministic in nature; that is, if the outcome occurs, this is the way things will be. The determined state may have a variance as a result of the confidence attached to the estimate of the state; however, every attempt should be made to estimate the state independent of the probability attached to the occurrence of that state. One view that can be taken of the state structured for any one outcome is that it repre- sents the plan of action if it appears that external factors will cause that state to result.

Cost model inputs are developed from the variable states outlined. A cost model run is made for each outcome using the deltas from the reference state (reject outcome). If the innovation impacts several domains (e.g., training, job, etc.) then more than one model must be run. The costs for each outcome are summed to derive the overall cost (or savings) attributable to that outcome. The outcome states are those shown in Figures II-1 and II-2.

The strategy reflected by this step is that if the innovation can be justified on displaceable dollar costs alone, and without severely threatening liabilities, then any non-costable benefits become a bonus. In other words, any pattern of information that is in itself sufficient for making a decision choice "to accept" is by definition adequate information for that decision. Furthermore, most administrators in our culture are understandably more comfortable with demonstra- tions of merit based on financial advantage than on other variables.

Furthermore, the magnitudes of dollar investment, dollar return and risk give a picture of general scale of a potential commitment. If this scale is greater than the decision maker's budgetary constraints, further consideration about the proposed innovation is fruitless, no matter what its benefits may be.

There is also heuristic value in seeking to indentify costable elements. Variables that initially seem qualitative may be exploded into quantitative component variables at least some of which are costable. Where a decision involves a large commitment, considerable effort in identifying costable benefits is justified. Thus, some of the costable components in "level of morale" -- a nebulous concept

in itself -- is associated with absenteeism, psychosomatic ailments, accidents, readiness to work overtime, willingness to work under annoying or stressful conditions (and other factors). At least some of phenomena can be measured by frequency of occurrence and unit dollar costs. Such data can be used in estimating expected dollar benefits or liabilities. The combined contributions of such data may be substantial in projecting dollar outcomes in decision alternatives.

A parenthetical warning, however: Assume that the benefit for an innovation such as "moderately increased student morale" has been proposed, and its dollar benefits have been parsed out of this variable and entered into the sum of other dollar benefits for the innovation. If the benefit "moderately increased morale" is still retained and scaled as a utility, it is essential that it now be evaluated as the benefit of "moderately increased morale other than dollar benefits." If this exclusion is not made, the dollar value of the morale change may enter twice in the assessment operation. The result would be an inflated evaluation of the gross benefit picture.

On the other hand, if there are contributions in addition to the costable attributes that have been identified for the increase in morale, ignoring these factors would deflate the gross benefit picture developed about the innovation. It seems to be a normal human tendency to assume that where the greatest effort has gone, there lies the greatest value. If the major effort in analysis has gone into cost-benefit analysis, one can expect that non-costable factors will spontaneously tend to recede in psychological importance. This phenomenom has been observed in staff studies in business environments. There seems no clear remedy except the suggestion that at least some of the protagonists for an innovation remain separate from cost analysis so that their perspectives will not be blurred by participating in the hard work in that analysis.

Appendix I to this report contains some incomplete lists of operational components that may be associated with each of several frequently used qualitative variables. Most of these operational components can be counted as present or absent, and many may have unit costs associated with them. It is emphasized that these lists do not exhaust the possibilities of operational variables that may be associated with a qualitative or apparently non-costable variable. Nor does it mean that the qualitative variable is completely or exhaustively specified by a list of this kind.

Once, therefore, costable variables have been defined at each outcome, and input cases to models prepared, a total cost/savings can be attached to each outcome. The remaining qualitative variables will require specification in terms of direction from the baseline or present state, and in terms of magnitude as related to some operational context. The following step will prompt a more explicit definition of the operational meaning attached to qualitative variables.

Step 4 - PREPARE DESCRIPTIVE (GRAPHICAL) REPRESENTATION OF ATTRIBUTES

This step more or less extends the characterization of the outcome states resulting from the scenarios created in Step 3. Its purpose is to prompt a continued orientation to the operational meaning inherent in the selected relevant attributes. If this step does not cause the qualitative variables to be anchored in some operational context, then in all probability the variables should be eliminated from consideration.

It is important, at this point, that the assessor have a sound understanding of the goals of the organization into which the proposed innovation would be introduced. For each attribute impacted by the innovation, the assessor should be able to identify some organizational goal to which it can be related. The prior emphasis upon an operational and organizational context for each qualitative attribute should now be apparent. The assessor will attempt to describe the change to or away from some goal-state which occurs from introducing the innovation. Estimates will then be made for the states of each of the other potential outcomes. Figure II-4 provides a sample result of the estimation process. Attribute x_1 (On-Job Flexibility) is estimated at its present state (x_1^t) which is the state assuming the innovation is rejected. The assessor believes that this present state is about halfway (50%) between what is considered barely tolerable and what is considered ideal. Next, the assessor considers the effect of the innovation on this variable, if the innovation is successful (Outcome A). This situation is estimated to improve the state of the On-Job Flexibility from its present state (50%) to a point about halfway to the ideal state. This is shown as $X_1^·$ which is at the 75% point. Intermediate outcomes (B and C) are then estimated. This same process is performed for the remaining attributes (x_2, x_3 and x_4).

The preceding approach called for direct estimates of the outcome points by attribute. There are several other techniques which may be applied both here, and in a later step dealing with importance estimates. One or more of these alternate methods should be considered as a test for consistency (reliability) of the values estimated.

1. Ratio tests over range x_1^t to x_1^g. Considering the x_1 attribute shown in Figure II-4, test the reasonability of the ratios of each outcome level to the goal and tolerable states. For example:

   a) Is the present state (D) just as far away from the ideal as it is from the least tolerable state.

   b) Does satisfactory introduction of the innovation (Outcome A) accomplish 50 percent of what is expected in order to achieve the goal state (x_1^g).

   c) Does a failure to implement (Outcome C) result in essentially no change.

   d) If the innovation is implemented but fails to improve the system as anticipated, (Outcome B), does it contribute 50% as much toward accomplishment of the goal as successful introduction (Outcome A).

2. Ratio tests over range of possible outcomes. This is similar to the preceding test, except that the ratio of one outcome to another, independent of its relationship to the extremes, is checked. For example, consider the outcomes for the attribute x_2 in Figure II-4.

   a) Does successful introduction of the innovation (Outcome A) provide as much satisfaction or utility, as the unsuccessful introduction (Outcome B) reflects in decreased satisfaction or utility.
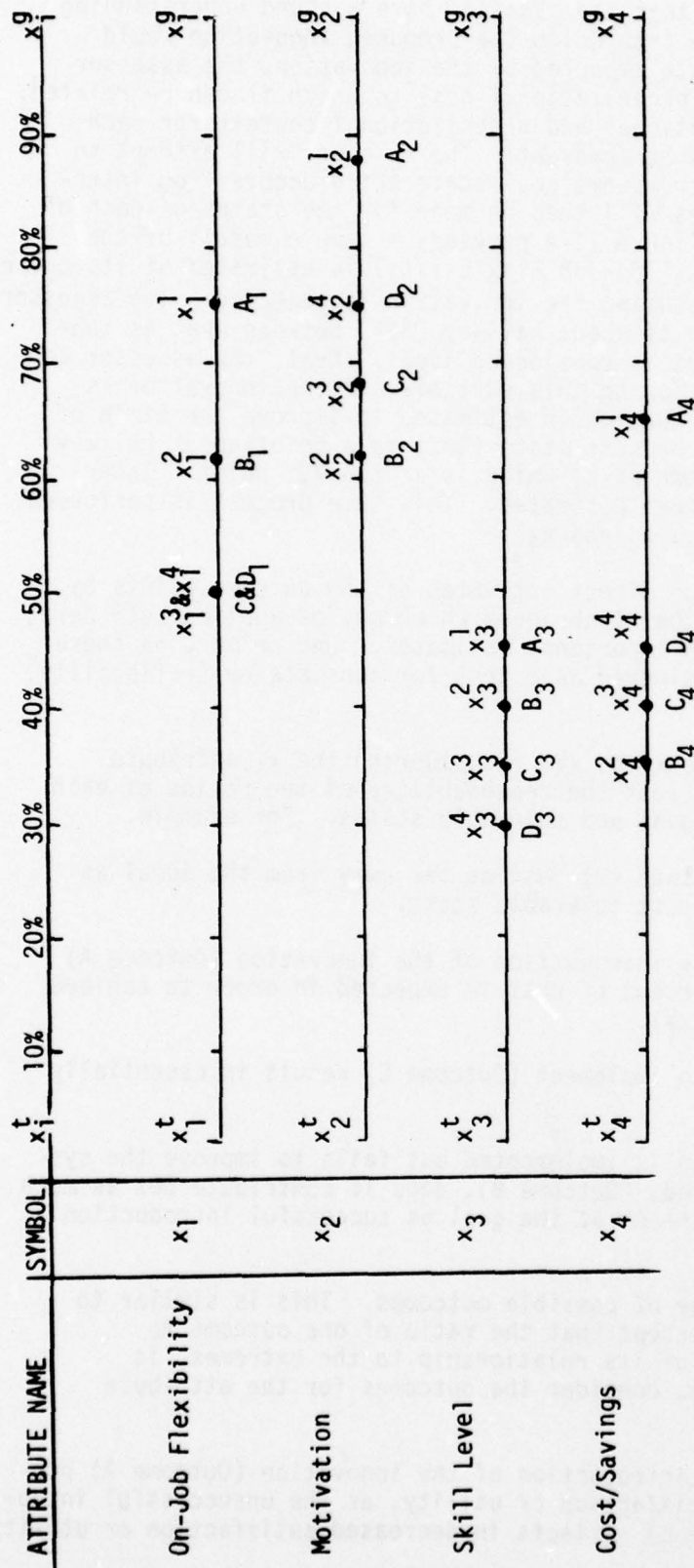
FIGURE II-4. GRAPHICAL REPRESENTATION OF OUTCOME STATES BY ATTRIBUTE IN RELATION TO LEAST TOLERABLE AND GOAL STATES

        b)   Does twice as much decreased satisfaction or utility result from unsuccessful introduction (Outcome B) as does from a failure to implement (Outcome C).

   3.   Indifference-lotteries over range x_1^t to x_1^g. The indifference-lottery method look for the point at which the assessor is indifferent between accepting the results of a specific outcome with certainty and taking a gamble at prescribed odds with the potential for receiving either one of two outcomes. For example, again considering attribute x_1 in Figure II-4, would you as the assessor be indifferent to maintaining the present state (Outcome D) with certainty, as opposed to taking a gamble with 50:50 odds of reaching the ideal state (x_1^g) or dropping to the least tolerable state (x_1^t). If this is reasonable, then the placement of Outcome D at the 50% level appears correct.

Step 5 - CONVERT ATTRIBUTES INTO UNITS OF A CONTINUOUS BASE ATTRIBUTE

If there exists one relatively important and quantified benefit variable, such as dollars, then this path through Steps 5 and 6 will be followed. Each level of each of the other attributes can be converted into units of the base attribute by applying a tradeoff procedure. The graphical representation of outcome states by attribute developed in Step 4 and presented in Figure II-4 will be used as the basic reference for this conversion process.

For example, in Figure II-4, consider that the dollars (cost/savings variable x_4) saved as a result of successfully introducing an innovation are a dominant aspect of the decision-making process. A specific dollar cost or savings will have been obtained as a result of the cost model processing of the outcome scenario parameters created in Step 3. Assume Outcome A produces a savings of $500,000 (500K) and Outcomes C and B result in losses or costs of 100K and 200K, respectively. Outcome D where the innovation has been rejected is neutral, or zero cost/savings. The dollar variable, then, has a relatively wide range of outcome possibilities, and is also significant in terms of the values taken on across the potential outcomes. The dollar dimension is therefore selected as the continuous base attribute into which the levels or states of each of the other attributes will be converted.

Consider the attribute, On-Job Flexibility, (x_1 in Figure II-4). Outcome A appears to have a substantial positive value in that it has the effect of moving the organization half-way to its ideal or goal-state (from a present state of 50 percent, to 75 percent). The question can be asked: "How many dollars of the organization's budget would the decision maker be willing to invest to avoid losing about the same level of On-Job Flexibility as results from successful introduction of the innovation?" The question should be answered independent of the range of cost/savings possibilities related to the innovation decision. If the answer were 50K dollars, then this represents additional value (over and above the 500K savings) to Outcome A. This same process is repeated for the remaining states of variable x_1, as well as for each state of the variables x_2 and x_3. Where an outcome is less than the neutral state (as with Outcomes B and C for x_2), then the appropriate question to be asked by the assessor would be: "How many dollars of the organization's budget would a decision maker be willing to pay to avoid the decreased level of the attribute resulting from a decision (and potential outcome from the decision) regarding the innovation?" The results of such a tradeoff analysis is tabulated, for example, as follows:

ATTRIBUTE

		Dollars(x_4)	Skill Level(x_3)	Motivation(x_2)	On-Job Flexibility(x_1)
O	A	+500K	+30K	+10K	+20K
U					
T	B	-200K	+20K	-10K	+10K
C					
O	C	-100K	+10K	-5K	0
M					
E	D	0	0	0	0

Intermediary verbal descriptions for each level would undoubtedly be used in arriving at these types of value estimates (e.g., significant increase in skill level). These should come about through the analysis from Step 4.

Step 6 - TRANSFORM BASE ATTRIBUTE VALUES INTO SYSTEM UTILITIES

As a result of the previous step, each level or state of all relevent attributes was converted into a value reflected in the same units as those of a selected important continuous base variable such as dollars. If the organization or system utility for units of the base dimension is linear, i.e., 200K dollars is preferred exactly twice as much as 100K dollars, then the values identified for each outcome state can be summed across Outcome A, then Outcome B, etc. For example, the value of Outcome A would be the sum of 500K dollars, 30K dollars worth of skill level, 10K dollars worth of motivation, and 20K dollars worth of on-job flexibility; it would be 560K dollars.

In many cases the organization or system utility for dollars will not be linear. The initial 100K dollar savings may be worth much more than another 100K (200K total) because it has greater utility in any application the organization has for dollars. This should be clear if the first 100K could be invested in a machine that would double capacity, increase profits, and produce a 100 percent return-on-investment; while the second 100K has a utility ten times that of the second 100K. This may be an extreme example, but it should make the point.

The problem, therefore, in this step is to insure that the utility ultimately determined at some particular outcome reflects the sum of the utilities for each of the attributes estimated at this state. An important assumption in this process is that when a tradeoff estimate is made for some attribute state in units of the continuous base variable, not in terms of its actual value. For example, when the system through the action of the decision maker is willing to commit (as is expected when the tradeoff value is decided upon) 30K dollars to achieve some X units worth of increased skill level, that committed amount takes into account the utility that 30K dollars has to the system. Therefore the 10K dollars estimated for an increased level of motivation would have one-third the utility of the 30K only if the system preference for each 10K increment were equal. Thus, before the utility of an outcome can be determined, a utility function must be assessed across values of the base attributes. Procedurally this can be accomplished as follows:

1. Specify the most and least desirable states of the base attribute. For example, the most desirable may be +500K dollars while the least desirable may be -200K dollars. Let these states be assigned arbitrary utilities of 100 and 0, respectively.

2. Select intermediate points and using a direct estimation, ratio; and/or lottery technique, determine their utility on the 0 to 100 scale. These techniques were presented in Step 4.

3. Graph the utility function for the base variable in order that states of the other attributes can be transformed into a corresponding utility figure through interpolation. Once the utility function has been graphed, it seems more psychologically sound to transform the utility scale by making zero utility equal zero units of the base variable, thus providing both positive and negative values for utility.

Figure II-5 shows an example of a graph for the utility of dollars. Some of the interpretations made from it are as follows. They also represent samples of the means by which the graph could have originally been contructed.

a. There is as much utility to the system in obtaining 500K dollars as there is disutility in losing 200K dollars.

b. The system represented by the decision maker is relatively indifferent to a 50:50 gamble between winning 500K dollars and losing - 200K dollars. The decision maker is, therefore, conservative and exhibits a substantial aversion toward risk since a risk neutral decision maker would accept the same gamble when the chances of winning the 500K dollars are only about 30 percent. $((p(500) + (1-p)(-200))$ p = 28.6%).

c. There is as much utility to the system in receiving the 100K dollar increment from 200K to 300K as there is in receiving the 200K dollar increment from 300K to 500K. In other words there appears to be diminishing marginal utility for each additional 100K dollar increment.

d. There is four times the disutility from losing the first 100K dollars as there is in losing the second 100K dollar increment. (Perhaps losing 100K is so damaging that the additional 100K cannot do that much more harm.)

The states of all attributes previously converted into units of the base dimension can now be translated directly into utilities by interpolating from the graph. The following example taken from Step 6 shows the results. Note that the utility scale having both positive and negative units is used. The original dollar values are shown in parentheses.
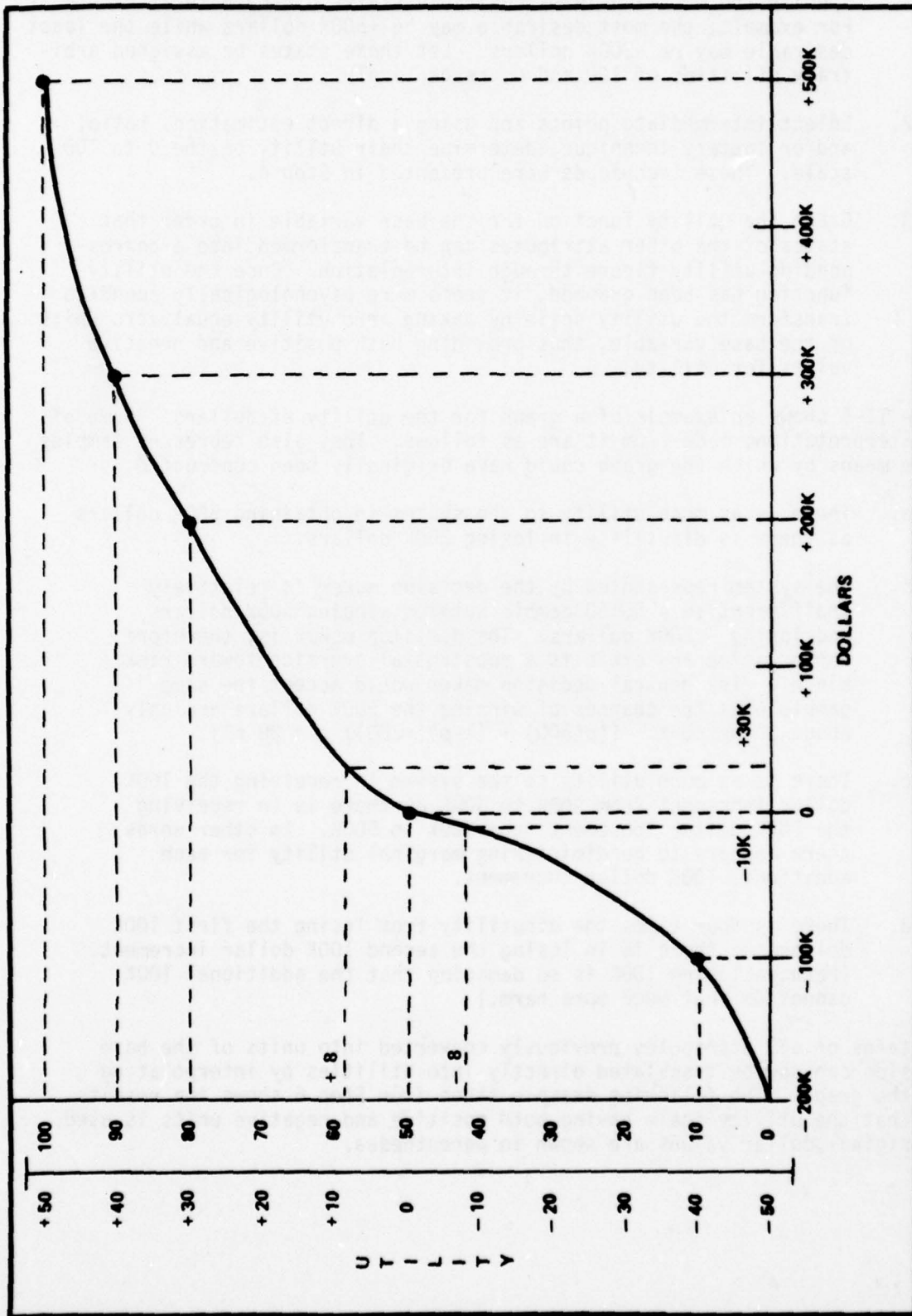
FIGURE II-5. GRAPHIC EXAMPLE OF A SYSTEM UTILITY FUNCTION OVER THE ATTRIBUTE , DOLLARS

ATTRIBUTE

		Dollars(x_4)	Skill Level(x_3)	Motivation(x_2)	On-Job Flexibility(x_1)
O	A	50 (500K)	8 (30K)	2 (10K) Est.	5 (20K) Est.
U					
T	B	-50 (-200K)	5 (20K) Est.	-8 (-10K)	2 (10K) Est.
C					
O	C	-40 (-100K)	2 (10K) Est.	-5 (-5K) Est.	0 (0K)
M					
E	D	0 (0K)	0 (0K)	0 (0K)	0 (0K)

The total utility for each outcome can be determined by adding across:
Outcome A = 65, Outcome B = -51, Outcome C = -43, and Outcome D = 0.
These utilities will be checked for consistency in Step 9 and the results
incorporated into the decision tree in Step 10.

Step 7 - ASSIGN IMPORTANCE LEVEL TO EACH ATTRIBUTE

The following technique will be applied when it has been determined that
no dominant continuous base attribute exists. This and the following Step
8 will permit the development of a system utility figure for each decision
tree outcome; thereby parelleling the objective of Steps 5 and 6 where there
exists a dominant base variable.

1. Consider the goal-related operational effects projected to be
    achieved from successful introduction of the innovation (points
    A_1, A_2, A_3, and A_4 in Figure II-4.

2. Select the most important effect considering the importance of the
    system goal (all points X_j^g in Figure II-4), the present status of
    the attribute, and the size of the incremental change resulting
    from successfully introducing the innovation into the system.

3. Select the least important effect considering the same factors as
    in 2 above. These are the points designated as X_i^t in Figure II-4.

4. Rank all other attributes considering the same factors as in 2.

5. Assign the most important attribute selected in 2. a value of 1.0.
    Directly estimate the importance of other attributes in relation
    to the most important one.

6. Check the consistency of the importance weights assigned by the
    direct estimate technique by performing one or more of the following
    additional techniques for estimating importance.

    a) Assess ratios of importance between all combinations

    b) Assess points of indifference using a lottery technique.

Returning to the example presented in Figure II-4, assume that the continuous base variable costs/savings (x_4) is relatively unimportant, and that it is felt more appropriate to use the procedures in Steps 7 and 8 for assessing the system utility at each outcome. The following sequence of thought-processes might illustrate the technique for arriving at relative importance values for the four attributes of Figure II-4.

    a.   The most important attribute is Skill Level (x_3) since it is an important system goal and its present state is relatively low.

    b.   Motivation is least important since it is presently at a fairly high level.

    c.   On-Job Flexibility is considered more important than Cost/Savings because it contributes significantly to overall mission in terms of preparedness.

    d.   Skill Level is assigned a value of 1.0; the remainder in order are On-Job Flexibility, Cost/Savings, then Motivation.

    e.   The Motivation attribute is less than half as important as Skill Level; it is therefore assigned an initial value of .4.

    f.   On-Job Flexibility is nearly as important as Skill Level. Motivation is about half as important as On-Job Flexibility, therefore On-Job Flexibility is assigned a value of .8.

    g.   The Cost/Savings variable is slightly more important than Motivation, and is judged about half as important as Skill Level; it is, therefore assigned a value of .5.

The final importance weights to be assigned each of the four attributes will be:

Skill Level	- 1.0
On-Job Flexibility	- .8
Cost/Savings	- .5
Motivation	- .4

These will be applied to the scales shown by the example in Figure II-4 to develop a system utility by outcome.

Step 8 - TRANSFORM WEIGHTED ATTRIBUTES INTO SYSTEM UTILITIES

At the conclusion of Step 7 each benefit variable has been quantified in terms of contribution to some specific system goal, and its weight of relative importance to the overal system mission has been estimated. In this step, the relative importance weighting factor will be applied to adjust the lengths for each attribute; the result will be a corresponding

assignment of system utility to each outcome of the rescaled attributes. The following technique is applied to determine the utility for each outcome. Figure II-6 presents an example of the results.

1. Assign an arbitrary scale for system utilities with zero as the midpoint, the midpoint representing the present system state.

2. Select the scaled attribute weighted as most important and anchor its present state (point D) at the zero point of the arbitrary scale.

3. Let the total length of the most important attribute scan exactly 50 percent of the arbitrary system utilities scale (this is for convenience and will permit all other attributes of lesser importance to be contained within the scale). Skill Level (x_3) in Figure II-6 is shown to extend from zero to plus fifty since its anchor point (D) was at the left most point on its scale in Figure II-4.

4. Select the next most important attribute and anchor its present state (point D) to the zero point of the arbitrary scale.

5. Let the total length be proportional to the length of the most important attribute based upon its relative importance weight as determined in Step 7. In Figure II-6, On-Job Flexibility with a relative importance of .8 is shown to extend a length of forty system utility units (.8 x 50). Since its anchor point (D) is also at the left most point on its scale from Figure II-4, it is shown to extend from zero to plus forty in Figure II-6.

6. Repeat 4. and 5. above for the remaining attributes. Figure II-4 shows that the other two attributes, Cost/Savings and Motivation, extend over twenty-five (.5 x 50) and twenty (.4 x 50) units, respectively. Since their anchor points (D) are between other outcome points, their scales extend in both a positive and negative direction from the zero or midpoint.

7. Read the system utilities for all A outcomes from the scale and add them algebraically, Thus from Figure II-6, $x_2^1 = 10$, $x_4^1 = 17$, $x_1^1 = 40$, and $x_3^1 = 50$; the total of all A outcomes is 117.

8. Repeat 7. above for Outcomes B, C and D. Outcome D will always be zero. Thus from Figure II-6, $x_2^2 = -10$, $x_4^2 = -8$. $x_1^2 = 20$, and $x_3^2 = 34$; Outcome B total is 36. Outcome C totals 8 with $x_2^3 = -5$, $x_4^3 = -4$, $x_1^3 = 0$ and $x_3^3 = 17$.
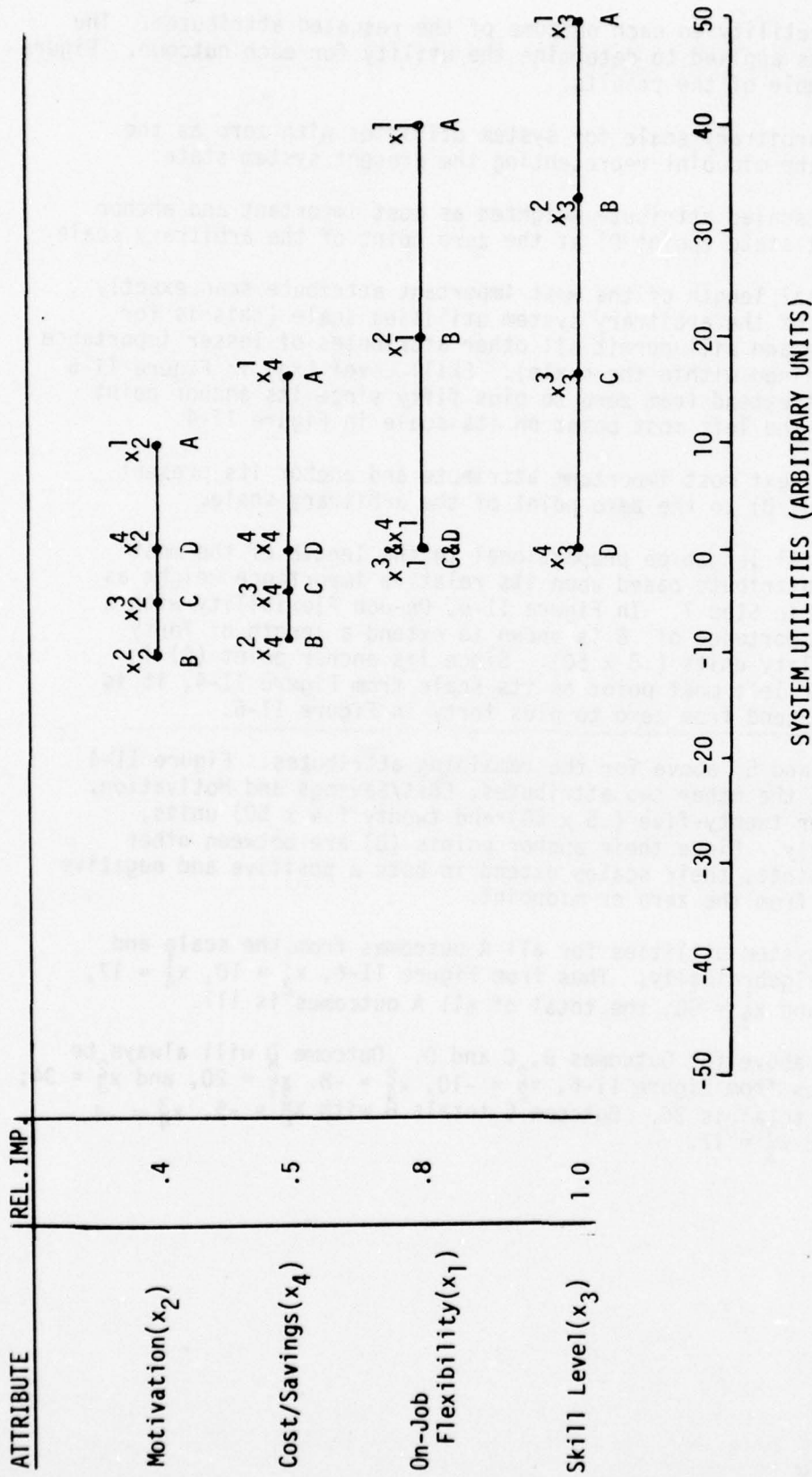
FIGURE II-6. APPLICATION OF IMPORTANCE WEIGHTS TO DEVELOP OVERALL SYSTEM UTILITY BY OUTCOME

In summary, then, the total of the relevant attributes across each major outcome is as follows:

ATTRIBUTE	OUTCOME A	OUTCOME B	OUTCOME C	OUTCOME D
Motivation(x_2)	10	-10	-5	0
Cost/Savings(x_4)	17	-8	-4	0
On-Job Flexibility(x_1)	40	20	0	0
Skill Level(x_3)	50	34	17	0
Total	117	36	8	0

These are the utility figures which will be inserted into the decision tree in Step 10 after a consistency check has been performed in Step 9.

Step 9 - PERFORM CONSISTENCY CHECKS OF OUTCOMES AND OF PROCESS

At this point in the procedure, some system utility for each decision tree outcome will have been generated, either as a result of the conversion of relevant variables into units of a continuous base variable as was performed in Steps 5 and 6, or through application of the techniques outlined in Steps 7 and 8. Both Steps 6 and 8 recommended crosschecks for the consistency of the subjective estimates. In this step, a wholistic common sense appraisal of the relative total utility of each outcome is proposed. The consistency check steps are as follows:

1. Check the ranking of the outcomes for reasonability. From the examples summarized at the end of Steps 6 and 8, the following total utilities by outcome were obtained.

	Step 6	Step 8
OUTCOME A	65	117
OUTCOME B	-51	36
OUTCOME C	43	8
OUTCOME D	1	0

    As is seen the results are quite different between the two, however, the assumptions were quite different and there was no intent that the results in any way be similar.

    At this stage, the assessor should make the intuitive assessment on the order of outcomes for each of the innovations leading to the above summarized results.

a) Is the order reasonable for the innovation processed through Step 6, i.e., A, D, C, then B?

b) Is the order reasonable for the innovation processed through Step 8, i.e., A, B, C, then D? Does it seem reasonable that acceptance of the innovation, no matter what the final outcome, will provide more utility to the system than outright rejection?

2. Check the ratios of outcomes to one another for reasonableness.

a) From the Step 6 results, does Outcome B appear to provide as much negative utility as Outcome A provides positive utility? Are Outcomes B and C just about as bad for the system?

b) From the Step 8 results, does Outcome A appear to provide over three times the utility as Outcome B?

3. If the results of 1. and 2. above do not appear to be reasonable in terms of the separations of the values of the outcomes, and their negative or positive direction with respect to the present system state, reexamine the following.

a) Have all relevant variables (impacts of the innovation) been explicitly identified and/or taken into account?

b) Have realistic scenarios been created for each outcome considering operational effects?

c) Have the individual attributes been realistically assessed with respect to the system's goals?

d) Have the attributes been weighted realistically?

4. The need to perform extensive reconciliation analysis in 3. above should be based upon a sensitivity analysis of the decision variable after the decision variable has been calculated taking into account the probabilities within the decision tree.

Step 10 - FINALIZE DECISION TREE AND COMPUTE DECISION VARIABLE

This step is unchanged from that outlined in Task 04.05 of the ETAM Procedures outlined in the Phase II-A Final Report. The decision tree into which the estimated utilities will be inserted was shown in Figure I-3 of this report.

The techniques which have been outlined for the subjective estimation of utilities are also appropriate for probability estimates to be incorporated at junctures W, X, W1, etc. in the decision tree diagram of Figure I-3. Section III of this report contains a good deal of the rationale as well as caveats that the estimator should be aware of in making probability estimates. Also, reference is made to a number of the forms of bias which may cause discrepancies in estimates.

As with utility estimating, it is extremely important that the assessor identify the operational context used as the basis for the probability estimates. Task 3 of the ETAM Procedures (Phase II-A Final Report) identified the important factors relative to probability estimates of the success or failure of an innovation. Some of these factors were organizational compatability, technological state-of-the-art, technical support, attributional acceptance, etc.

The other major input to Step 10 will be the risk reduction packages originally outlined in Tasks 3, 4, and 6 of the ETAM Procedures. These packages will affect the probability values in the decision tree diagram (Figure I-3).

The final stage of this step is to fold back the decision tree in order to arrive at a utility figure for each alternative path. The alternative path with the highest expected utility would normally be the one recommended to be followed by the decision maker. This of course assumes that the decision maker has the "maximization of expected utility" as the operating decision principle.

Step 11 - PERFORM SENSITIVITY ANALYSIS

Task 06.07 of the ETAM Procedures outlined the technique for handling the sensitivity analysis step. A key to the performance of this step is the assessor's estimates of confidence in previously estimated factors, such as probability, utility, and relative importance. The expression of confidence is in the form of a most optimistic and most pessimistic estimate of the value that the variable could be expected to take on. The sensitivity analysis tests the decision variable across this range to determine if there could be a change in the recommended decision. If the recommended decision alternative is the same, no further analysis of the variable is required. If the decision changes within the range (optimistic to pessimistic), then a probability that the change could actually occur is calculated. The ETAM Procedures should be referred to for additional detail on this process and its related assumptions.

Step 12 - PREPARE INPUT TO THE DECISION MAKER

ETAM Procedures Task 8 described the series of subtasks for framing the assessment data for presentation to the decision maker. It consisted of:

1. Assembling the information generated in each of the ETAM tasks. In the context of this report, the information to be assembled would be:

   a) The relevant variables from Step 1.

   b) The outcome scenarios from Step 3.

   c) The graphical descriptions of the impact of the innovation on system goals from Step 4.

   d) The results of the assignment of system utilities either from Steps 5 and 6 or from Steps 7 and 8.

e) The results of consistency checks from Step 9.

f) The folded back decision tree diagrams from Step 10.

g) The results of the sensitivity analysis which was performed in Step 11.

2. Defining the presentation format both in terms of style and time requirements.

3. Preparing the material in a format suitable for presentation to the decision maker.

4. Reviewing the material in terms of the objectives, assumptions, and alternatives.

Step 13 - MAKE DECISION

The entire ETAM Procedure including the previous twelve steps for refining the estimation and scaling process were performed so that the decision maker would have an appropriate information pattern for making a sound decision about the innovation. Based upon the value of the decision variable calculated for each alternative decision path, a reject, an accept, or an accept with risk reduction projects could be recommended. The decision tree structure provides, also, for multiple groupings of risk reduction packages to be analyzed. Therefore, it should be possible for the assesor to have formulated and analyzed a number of packages in order to show a degree of optimality in the final recommendation.

It is obvious that no matter how meticulous the assessor is in carrying out the assessment process, there are certain indisputable requirements.

a. The utility determination process must be conducted within the context of the goals of the organization into which the proposed innovation is to be installed. The assessor must relate to the same utility system as the decision maker.

b. Information acquired in the process must be reasonably accurate, either based upon a historical analysis which forms a reasonable basis for future projection, or an estimate by persons having recognized expertise in the area about which the information is being obtained.

c. The combinatorial logic for determining the value of the decision variable should be in reasonable agreement with that of the decision maker. This includes agreement in terms of the decision principle, "maximization of expected utility."

d. Finally, the assessment process should reflect a high degree of diligence on the part of the assessment team in acquiring and processing data.

The ETAM process attempts to attach utility to various decision alternatives as related to the decision about a single innovation. It does not preclude a comparative assessment between alternative innovations intended to accomplish a similar objective. It does not, however, extend to the decision process relative to which investments are most appropriate for an organization to make with its available funding. Techniques, however, have been suggested in the literature for application of decision and utility theory to the selection of investment type projects. The reader is urged to refer to the Bibliography at the end of this report for more extensive literature references.

SECTION III

ORGANIZATIONAL AND BEHAVIORAL RATIONALES FOR SCALING METHODS

INTRODUCTION TO THE RATIONALES

In Section II of this report, a step-by-step outline of a scaling procedure for range-of-effect assessment was presented. This procedure was developed from the rationales and references that are discussed in the following pages. Further developments in this project will enable these procedures to be used at an interactive display terminal. A number of functional elements and some information content will have to be added in order to enable flexible and useful three-way dialogue between machine system, the assessor and the decision making executive. It seems essential, therefore, to present not only the reasons and purposes behind the procedural structure but some substance to assist in its operational use. The content in this section will be reshaped into an "operating guide" for the intended users of the scaling and decision techniques.

Part A in this Section is primarily a justification for the use of reductionist techniques to supplement, if not replace, global and intuitional decision formats. Utility theories and decision models, a subset of which are being used, have not gained universal acceptance by any means, either on logical or pshychological grounds. It therefore seems important to clarify the justification for using them in ETAM. Decisions about proposed innovations are apt to involve large commitments if they are accepted, and sometimes equally large liabilities if they are not; and, because of their intrinsic nature, decisions about innovations have a larger context of uncertainty, guessing and risk-taking than most decisions about plans and operations in an establishment. Offering procedures that claim to make good or better-than-good decisions in this context therefore deserve careful scrutiny in order to assure that it is not only efficient but wise to use them. In Part A, a kind of information "task analysis" of the decision-maker's problem is made. Considerable attention is given to the various kinds of reference values that the executive might or should use -- and share with his assessing staff -- in formulating decision policies. Structural decisions, as contrasted with operating control decisions, deal with relatively permanent institutional changes, and are therefore applicable to the innovation problem.

In Part B of this Section the rationales for reductionist formats for decision making are continued but in a more specific, behavioral context. Liabilities in human processing of global decisions are discussed, while still accepting the potential validity of this method at least as collaborative with reductionist methods. Various liabilities and assumptions in the latter methods are also explored -- with the help of the literature -- and practical conclusions are drawn. The liabilities and limitations should serve as guidance for things to avoid, or at least to counter, in the design of semiautomatic scaling procedures. This body of content will also yield training material as well as system design guidance.

PART A. STRUCTURAL DECISIONS ABOUT LONG-TERM ORGANIZATION MISSIONS

TWO TYPES OF DECISION ANALYSIS

A large amount of methodological research during the past several decades has centered around human decision processes. The literature is known variously as "decision theory", "behavioral utility theory", and others. As is to be expected, there are some who vigorously accept the methodology, some who are skeptical, and others who are hostile.

Those who are hostile would probably say that any human decision must be based on a global context of information. This information consists of known data, expectations, resource commitment and a scheme of personal and organizational values. The processing of those data into a decision is a complex and delicate operation and dynamics of which will not survive the surgery of analysis or the constraints of artificial structuring imposed from the outside. These individuals believe in what will be called the wholistic and intuitional mode of decision making. The present authors respect this point of view, but not exclusively.

The skeptic who may be informed about the literature will point out that clearcut operational evidence for the superiority of the analytic techniques of formal decision and utility theory has not been developed (at least for non-diagnostic types of decision making). And the indirect evidence based on behavior studies in artificial decision environments, or about limited kinds of decision making tasks, or on psychologizing about "information overload" is not necessarily convincing. The skeptics have a point, too, with which to be sympathetic. In their moments of candor, some of the researchers themselves acknowledge that the crucial empirical comparisons still remain to be made--if in fact such comparisons are really possible.

The protagonists of decision theory and utility have an impressive literature and the support of many convincing logical and psychological arguments for analytic and reductionist--or what will be called "decomposed"--techniques for the structuring of judgments about problem variables and relationships amoung problem variables. It is true that many researchers have as their objective an accounting for the human behavior in wholistic judgmental decisions by analytic "models" of that behavior. But the reciprocal concern is with methods to assist the decision maker in being rational and consistently rational in assessing evidence and applying value judgments to that evidence. The techniques are intended to aid the decision maker in using to the fullest extent all of the information he has that may be relevant--and this includes impressions, beliefs, values and judgmental processes. But they also are intended to help the decision maker uncover his various assumptions about his values and "utility functions" and thus to be consistent with them, or to modify them if need be. This point of view which, more than tentatively, uses formalism in structuring the decision-making process is espoused here.

It is possible to be consistent with both the wholistic and the decomposition or analytic doctrines because the prodedures proposed in the earlier Section II embrace both doctrines, so that the outcome of one serves as a check on the outcome of the other for selecting a decision alternative. This expedient is

not intended as a compromise between hostile points of view. Rather it is a recognition that the wholistic and the analytic procedures both have value, so that one mode should support the other. And, finally, the decision-maker himself should decide on which procedural options to use if he prefers one exclusively to the other.

There is another equally practical consideration. In most cases, it is realistic to expect that the assessor who collects, organizes and presents the data, and the considerations and conclusions about the decision problem, will be in a staff relationship to the decision-making executive. The assessor himself will present both wholistic and decomposed solutions to the executive decision maker. It is proper for the executive to view, like the successive layers of an onion, the pattern of constituent data, attributes, values and judgments -- including probabilities of various projected outcomes -- so that the executive may selectively intervene. He may apply somewhat different value systems, projections, and interpretations than those of his staff. He may wish to make selective and specific modifications to the decision "model" he is examining, perhaps at an interactive computer terminal, and observe the change in projected outcomes from various alternatives for the decision. Obviously, he could not interact this way with a wholistic evaluation.

A computerized model can take revised inputs and quickly recompute the reflection of these revisions into the overall evaluative output of decision alternatives and their projections of outcomes. This capability can be an important assist from a procedural model having a structure that is reflected in computerized operations.

The practical value of analytic procedures that is gained by selective interaction, either by the assessor or by the executive decision-maker with any stage of conclusions should, in itself, be a sufficient justification of decomposition methods for dealing with complex problems with risky outcomes.

CHARACTERISTICS OF DECISIONS ABOUT ORGANIZATIONAL MISSIONS

Let us now home in on decisions to be made about innovations. An innovation is defined as any structural change intended to improve the efficiency (resource cost) or effectiveness of the organization over a long-term run. As Peterson (1971) points out, such decisions should not be oriented to maximizing the expected utility for any individual (such as the decision maker) but to maximize the expected utility accruing to the organization and its mission. "Expected utility" is a combination of the worth of a magnitude of benefit and the probability of achieving the benefit, assuming that a given commitment of resource has been made. Even though the decision-making executive attempts to take an "objective" organizational and systems orientation, it is inevitable that some personal and subjective factors will influence his choices -- his perceptions, values and risk-taking characteristics as well as momentary pressures and moods.

Unlike traditional business establishments, the military training and operational systems do not necessarily have money as their dominant utility reference, except as, perhaps, a limiting condition. Rather, their dominant criterion of worth consists of the goodness of the pattern of services they perform. These services form a network of functions and functional criteria that transform human recruits in the form of input aptitudes, skill levels, and attitudes into output of technical and combat skills at levels which meet

military mission requirements that are continuously changing.

Executive decision making in the military is likely to be a heterogeneous process, performed by a variety of individuals with varying backgrounds and interests. Many will have developed their own intrinsic "formats" or procedural models for converting evidence into alternatives and decisions. Many can point to the success of decision outcomes in the past and justify hesitation and even resistance to changing their mode of making up their minds. They may emphasize the dangers of looking at procedural trees for gaining a proper perspective of the forest. We have not found any literature which describes how to make the social and psychological transition from the traditional methods of institutional and wholistic decision making to the analytic process at the level of highly placed, successful executives. Huber (1974), however, suggests that the consultant serve as change agent for the client.

There are a number of constraints that are properly placed on the introduction of procedures based on contemporary decision theory and utility theory into the practical world of executives and their staffs. The procedures should err on the side of simplicity rather than on the side of complexity and finesse. The simplicity of the procedure should enable the assessor (and executive) to keep in mind the information he is having to process at the same time that he can keep in mind what he is doing in each procedural step imposed on him. Thus he can apply "common sense" to the procedure, and readily make corrections in the context of the decision problem to solve, rather than merely the procedural issues to resolve. The problem-solving format should not obscure the real substance of the problem to be solved.

Fortunately the research literature supports simplicity in procedure. A dominant finding among the critiques of decision methodology is that simplicity in procedure and procedural assumptions is almost always just about as good as more complex and "sophisticated" procedures and assumptions. "Just about as good" means that it explains practically all of the variance (to within a few per cent) that is generated by the more sophisticated model.

If computer support is available, it may be tempting to introduce procedural niceties and complexities with the objectives of increased reliability and "precision"--even where the practical meaning of precision is unclear. Our explicit policy, therefore, is to maximize the understanding by the human of what is going on in the scaling task, both by simplifying it and by exposing it to his comprehension.

In the decision-making problem we are trying to solve there is another virtue to simplicity in scaling procedures. This virtue comes from the desirability of the assessor to communicate to the executive decision maker the operational assumptions, data, probability estimates and procedures that have gone into the picture presented to him. Thus the decision maker can selectively retrieve and review the work of the assessor at any of the many stages in the assessment operation. The rapid comprehension by the executive of the contextual meaning of these processes can readily be fogged by a maze of procedural detail.

Aside from its computational function, we see the major purposes of the interactive computer system as selective retrieval of data and data reductions, and for presenting the assessor with more or less equivalent procedural alternatives

for scaling operations. By performing the same task with procedural alternatives, the assessor may discover inconsistencies among outcomes and seek to reconcile them, presumably with the greater insight gained by each of the two or more methods.

The assessment task and its objectives may overlap, but are distinguishable from, the task of the executive decision-maker. The task of the assessor is primarily staff work: (1) data collection about the costs and benefits and liabilities of the proposed innovation for the training establishment across the full range of its potential effect, (2) applying his picture of probabilities and system values and needs to these data, and (3) reducing the entire mass of information to a presentation that the experienced executive decision-maker can assimilate, in terms of the presented alternatives, generally on the order of an hour. During this period, the executive may seek to determine whether the assessor's values and priorities are at least roughly similar to his own picture of the needs and values of the organization as a system. He may impose a different estimate of the probabilities of achieving various outcomes -- he may be aware of means of controlling the outcomes not taken into account or of dependencies not considered by the assessor.

The executive's choice of an alternative may often have two contexts. In one context, an absolute judgment about the innovation answers the question: "Is the commitment arising from some form of acceptance of the proposal within the discretionary limits of the executive's budget, taking into account also the risk and size of expected return?" This judgment is based on the merits and costs of the innovation solely on its own terms. But executives frequently have more than one proposal clamoring for support and these also demand dollar commitments. The second context for making a choice is therefore comparative: "Since only one or two proposals can fit into allowable budgets, which one is the best bet, all factors considered?" The second kind of choice may be more difficult than the first, and probably is representative of real conditions.

The executive may also feel assured that he has been given the most efficient way of achieving the benefit objective of the innovation if the ETAM Procedures Task 2 (Develop/Examine Alternatives to the Innovation -- see Figure I-2) has been performed.

A critical factor to underline here is the relatively small amount of time usually available to the executive to make any single decision, even an important decision involving a long-term commitment. This seems a practical consideration in the real world of time-stressed executives. The assessor as staff to the executive does not normally have the same stringency of time limitation, and certainly not in the same scale. After bringing together the data, it becomes his task to present information that is so reduced and organized that the executive can quickly perceive both the gross picture of overall worth and its structural components. The structural components consist of (a) the alternatives and their outcomes, (b) attributes and magnitudes of benefits and liabilities, (c) relative importance of each attribute, (d) probability of alternative outcomes, and (e) any estimates of the range within each benefit variable within which the outcome might fluctuate. (The last factor helps in making a sensitivity analysis.) Scaling techniques can generate a single figure of merit for all of these components together. They can also reveal contributory figures of merit that were developed for component

This realistic concern about the time-stressed executive may suggest one useful criterion for testing decision-making techniques: does it increase the effective bandwidth of the executive? This is, does it enable him to make a decision with the same confidence in understanding the problem after one hour interacting with it as contrasted with the many hours he would spend on it with alternative forms of presentation and interaction?

The specific matter of decisions about accepting or rejecting innovations has another major facet. It is a characteristic of institutionalized organizations to become conservative. Conservative policies and criteria are manifest in the principle of making no changes except those that have some assurance of demonstrably reducing displaceable costs in existing operations. This policy inhibits entrepreneurial decisions that are, by definition, risky and deal with substantial changes both in process and in objective. Thus, the ability to formalize the significance of values in addition to those of dollar costs, and into which dollar costs enter only as one dimension of the decision, may offset the erosive effects of exclusively conservative policies. The keynote here is in making explicit--as contrasted with implicit and intuitive--the value system behind the decision policy. The formal techniques of decision theory not only encourage but force this explicitness.

THE UTILITY CONCEPT OF "MISSION GOODNESS"

If a proposed innovation has a variety of benefit attributes, combining the worth of these several attributes into a single figure of merit implies that the assessor (and decision maker) has some underlying continuum of worth into which each of the benefits can be translated. Let's call this continuum a reference utility.

Utility theory has traditionally offered two kinds of reference utility. One common example references dollars and "dollar equivalent." The test question would be: "How many dollars would be equivalent to the value of this benefit?" or "How much would you pay to avoid having this negative benefit?" This reference utility is common to problems in business contexts.

A second traditional reference utility is that of personal "satisfaction and dissatisfaction." A sample test question would be: "Would you be twice as satisfied if you were given Benefit X as if you were to be given Benefit Y?" In the ultimate sense, perhaps, every decision made by a human incorporates a personalized set of values. In the more immediate and specific sense, the issues at stake are the beliefs by the executive and his staff as to what changes in the system and its mission result in more efficient or effective operation. Ultimately these beliefs will have personalized references, so that the "satisfaction-dissatisfaction" continuum has theoretical justification.

But a third kind of basic reference utility is proposed here, and will tentatively be called "system goodness" or "mission goodness." The reader is reminded that any utility continuum is an abstraction and simplification that is justified by its practical usefulness--as compared to alternative reference scales. So let's agree not to cavil about whether it is or is not a homogeneous continuum, just as traditional researchers and consultants have not been deterred about the homogeneity of the equivalent dollar utility or the satisfaction utility. Rather, let's examine some of the implications arising from the "system goodness" reference utility and see if a rationale can be developed for it.

A reference utility, whether based on dollars or satisfaction or system goodness, is likely to have a significant influence on the orientation of the assessor's and the executive's perceptions and judgments. Each of these references suggests a somewhat different role and a different set of conceptual tests of what information is relevant and how it should be slanted.

The system goodness utility depends for its usefulness on a number of conditions that are normally assumed to hold for the executive and his staff: knowledge of the variables that do actually operate in the system and the dynamics of these variables; this knowledge must inevitably be incomplete and partially erroneous, and because of the complexity of systems, sometimes inconsistent from one context to another.

The application of the system goodness utility depends on some consistent picture of the criteria of system performance. These criteria can have as their reference starting point some present set of operations and operational entities and networks. This base is similar to the reference status from which utility scales and functions are now developed in decision theory procedures. One evaluates a potential change and its meaning by comparing it with some present state of affairs. Thus one can assess limitations and liabilities between demand and response of the system now, extrapolate from present demands into the future and estimate the mismatch between (a) present system resources and operations and (b) those future demands.

Let us assume (or stipulate) that a competent executive and his staff are those who can distinguish between the operational needs that are current and the structural needs that are progressively developing over time into the future. The "future" may range from three to ten years, with perhaps a median point at five years hence. The executive's domain of responsibility may consist of an organization and an organizational mission, and its interfaces with other organizational missions. He may project past and present trends that increase the mismatch between mission goals and available resources such as people, skills and attitudes, technology, procedures, and the organization of these resources. He may simplify a staggering complex of variables in this mismatch into perhaps six to ten major attributes, functions or entities. For example, many executives in all sectors are viewing with increasing concern what is grossly called "the motivation of people to work." An individual executive and his colleagues may well develop a set of priorities or weights of importance to each of these attributes of extended concern. Assuming an environment where economics need not be completely restricted to here-and-now operational matters, practical attention can be directed toward one or a combination of ameliorative solutions.

We should ask the basis upon which the assessor and decision maker establish the relative importance of one benefit (or liability which can be treated as a "negative benefit") over other benefits that may be promised by an innovation. Simplistically, the answer would be: the most important benefit (defined as the combination of a variable and a given magnitude of change on that variable from a reference state) is the one that will have the greatest contribution to the efficiency and effectiveness of the training system and its mission. As it stands, this response is a circular statement like the assertion: "What is most important is what is most useful." The issue requires further clarification.

We must assume that the informed executive's utility structure, or scheme of

values, is a profile of weighted objectives for system behavior and accomplishment. This schema will be symbolized by dimensions in his conceptual image or model of the future system. In this conceptual model, means and ends may be relatively indistinguishable until an innovative proposal specifies a means with its prospective benefits as related ends. For example, one major dimension of problems in training systems in changing economic and social situations is that of the form of delivery of instruction. Another dimension may be perceived as that of geographic movement of students from operational environments to training environments. An innovation of a given kind may provide benefits in both of these dimensions, perhaps by eliminating one of the problems entirely but with the risk of creating new ones. Depending on the pattern of pressures which the executive foresees in the time period under consideration, and resources available without the innovation, his implicit model of system operations and values may give three times the weight to the elimination of student movement as he would give to the benefit in instructional function. These weights would enter into the assessment of the decision alternatives.

The decision to accept or reject the proposal would be based on the cost in resource (dollars) balanced against the extent to which the benefits of the innovation filled the void between his projection of system need or "requirement" on one or more dimensions of system operation, and the size of the remaining void after the benefits were in place. His confidence or estimate of subjective probability that the innovation would achieve the benefits would factor his perception of the worth of the promised benefit--but we are not concerned with the details of the subjective probability issue at this point.

The foregoing discussion of the factors in the executive's idealized perception of an innovation in the context of a system or organization can be summarized in a list. His conceptualization should consist of:

1. The present objectives of the system's mission in terms of options for the selection of its inputs, and the various criteria of the system's outputs at various interfaces.

2. The projected changes in these objectives expressed in the same terms.

3. Some set of gross, major variables that determine goodness of the system in operation. In training, where the output variables consist of quantity and quality of trained persons, the operating variables may be grossly characterized by the following list:

   a. Cost per student training hour, or cost per unit increase in a student's on-the-job capability.

   b. Minimum aptitude level for expecting the student to complete successfully given courses of instruction, or to achieve at lease a minimum level of on-the-job capability.

   c. Range and levels of on-the-job skills which formal training can instruct.

   d. Student hours to achieve a given criterion of mastery.

e. Attrition level in a course of instruction, holding aptitude and mastery level constant.

f. Flexibility in changing instructional capabilities to changes in input and/or to changes in output demand for type of skill, skill level, or range of skills.

g. Student motivation to learn and perform the jobs for which he is being trained.

And so forth.

Note that items a through e can have dollar costs computed for a given set of conditions. But improvements in any of the system variables can have benefits in addition to, or other than, dollar savings, and in many cases the utility of the non-dollar benefit may eclipse the utility of the dollar benefit.

Note also that although many or all of these variables may interact with each other in the process of selection and training, it is possible to introduce a change in one and let the others vary as a consequence of that change. Parenthetically, the understanding of the interactions among system variables is an important criterion for "knowing the dynamics of system behavior."

4. The status of variables identified in Item 3 at the present time in current operations of the system, taking into account present demands and present resources.

5. A projection of the future demands that are expected to be placed on the system and the capabilities of the present system to cope with them. "Capabilities" may be classified according to their operational relevance to the variables cited in Item 3.

6. A perception of the gaps between present demand-capability relationships and future demand-capability relationships.

7. A prioritization as to which of the projected gaps (and variable on which the gap appears) will be most important to close.

The rational executive would then tend to favor an innovation which promised to eliminate or reduce a gap on prioritized variables, and according to how much of the gap was closed. If the innovation appears to close only a small part of the expected gap of even a high priority variable, the executive may reject it because he knows that he will still have to find additional innovations for the same purpose. He may prefer a large magnitude of gap closure from a single commitment, and this is likely to be sound design policy for organizations where it is reasonable to expect substantial skrinkage of promised benefits.

The concept of system goodness that emerges is a composite of variables dealing with expected demands to be imposed on the system, projected capabilities of the system, and the mismatch perceived between patterns of demand and patterns of capability.

Figure III-1 represents a schematic of how the executive who is responsible for filling personnel quotas to a training school, and its quotas for various courses of instruction, might view an innovation in selection procedures that would give him "great flexibility in adjusting fluctuating supply coupled to fluctuating demand."
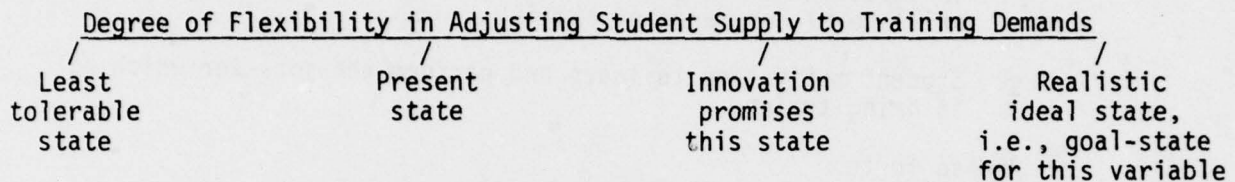
<u>Degree of Flexibility in Adjusting Student Supply to Training Demands</u>

/	/	/	/
Least tolerable state	Present state	Innovation promises this state	Realistic ideal state, i.e., goal-state for this variable

Figure III-1. THE CHANGE IN STATE OF AN OPERATIONAL VARIABLE
AS A CONSEQUENCE OF A PROPOSED INNOVATION.

From this executive's standpoint, the innovation, if realized, would shift the goodness of the "selection and assignment to training" procedure a substantial increment towards a realistic ideal. But this executive also has responsibilities for increasing the level of interest in potential recruits so that, ideally, practically all that are selected will want to take the training that is offered to them. In this domain of responsibility, the selection procedure problem is only one dimension of his "subsystem" problem. Figure III-2 shows the extent to which the innovation weights in his total problem.

Overall Subsystem Capabilities

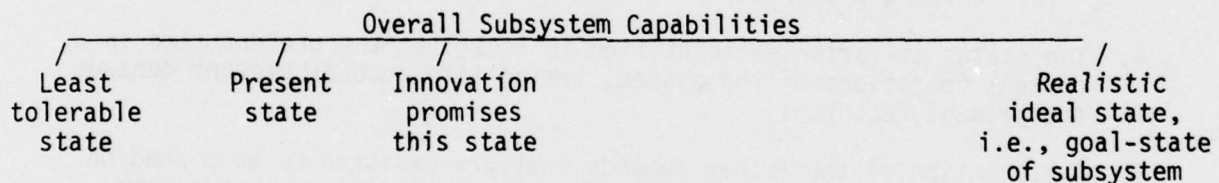| /	/	/	/
Least tolerable state	Present state	Innovation promises this state	Realistic ideal state, i.e., goal-state of subsystem

FIGURE III-2. THE CHANGE IN A SUBSYSTEM STATE OF BEHAVIOR AS A
RESULT OF INTRODUCING AN INNOVATION AFFECTING
ONE VARIABLE IN THE SUBSYSTEM.

Figure III-2, taken with Figure III-1, shows several things. One is that the present state of the <u>subsystem</u> is estimated to be lower (closer to the Least Tolerable State) than that in one component variable: the executive is having more trouble with motivating the selectees than in finding them. It also reveals that the contribution of the innovation to subsystem effectiveness is therefore scaled down from its magnitude when viewed only in the context of selection of trainees. This is a diagrammatic representation of how the execution might scale the relative importance of more than one benefit that might be provided by an innovation, as well as a schematic of how he might evaluate various innovations with respect to each other, assuming that the risks were equivalent.

The executive should also be clear about the time horizon which spans the limits of his planning. He should take into account trends in changing demands, changing resources and changing technologies and project their future from the present and its problems. Otherwise he is committed to

patchwork solutions as expedients to immediate crises. As an ideal policy, the introduction of innovations should somewhat lead the operational necessity for them. On the other hand, the utility of innovations, whatever their early levels of realization, inevitably wears out. This consideration should affect the evaluation of the benefits promised by a technological innovation. Will the magnitude of the benefit cope with the magnitude of the projected problem over, say, a five to seven-year projection? This issue of the time span affected by decisions recurs in differentiating operational decisions from design and structural change decisions.

As a realist, the seasoned executive is aware that any rational model of system variables and their operational interactions is too simple-minded to suffice for practical decision making. He also knows that there is a large proportion of randomness in real-life operations. The precision with which predictions of events, or the effects of introducing changes on events, is limited by the proportion of controlled events to random events in the system's behavior. The realist must be aware that while the threat of disaster rarely turns out to be as severe as anticipated, so also benefits rarely live up to their promise, even when the promise is bright. These regressive effects in system behavior can be attributable to a combination of randomness or slackness in the causal continuities in the system plus the tendency of human systems to maintain conservative stability by assimilating and nullifying perturbations. In this sense, the introduction of an innovation in an existing institutional system is a perturbation. A realistic executive may believe in the importance of an innovation's objective, and in its technical adequacy, but have little faith in its ability to survive the erosive effects not merely of organizational hostility, but of organizational indifference. To the extent that this formulation applies to a given executive, he should want his staff to differentiate the technical value of the proposed innovation from their estimates of probability of successful realization of the benefits, so that he can add his own "degradation factor" to their considered expectations.

DIFFERENCES BETWEEN OPERATING DECISIONS AND INNOVATING DECISIONS

Executives with successful backgrounds in making operating decisions can have difficulty in the evaluation and decision making associated with structural innovations for the system. Although innovations are made in operational contexts, and operations must be sustained through the introduction of innovations, there are differences that justify examination. In many cases, executives may be selected for the making of innovational decisions from the ranks of those who have made operating decisions.

By definition, an operating decision applies to keeping the existing system within a moving equilibrium. Operation deals with steering and control within the boundaries of an existing pattern of service demands, resource capabilities and operating rules and procedures. On occasion, a demand may become unusually great, or a facility may break down, or an alarming backlog develops, or the quality of an output service threatens to fall below acceptable limits. Control decisions may often require (1) a diagnosis of the correctable cause when presented with gross symptoms of inadequacy; (2) the identification of an available resource to deploy as a remedy; (3) the assignment of the resource; (4) a follow-up on the effectiveness of the remedy. The decision process in control deals typically with one-shot, transient situations. The primary objective is to maintain efficiency even through perturbations to the system

where efficiency is measured primarily in costs per unit of throughput service. This correction of a disturbance is limited to making a minimum change to the existing system that is still sufficient to reduce the disturbance.

The executive who develops through decision-making experiences the concept of a control model of a system must thereby tend to acquire somewhat conservative points of view about the introduction of organic change into the system's procedures, processing networks, resources, and service objectives. His experiences have concentrated essentially on preserving the integrity of the system as it is and as he knows it. He has learned to cope with the special kinds of uncertainties in system behavior with a set of working hypotheses about causes behind the facade of symptoms. He can cope with the uncertainties about what existing resources, human or otherwise, can do in containing a threat. Like the experienced automobile driver, he has learned the range of slop in system behavior and can compensate for it.

Although the procedural format for making control decisions may be similar to a format for making design decisions that change the system structurally, the conceptual model and values applicable to the latter may differ significantly from those applicable to the former. An innovation, according to the definition in the ETAM Phase II-A Final Report, is a "relatively constant or enduring change in the procedures, objects or functions used in any aspect of the instructional process which may be viewed as a benefit (or liability) and has associated costs". The executive making a decision about an innovation therefore has the role of a system designer. This role includes modifications to an existing design of a system--meaning here the training establishment and any of its parts or processes.

To the system designer, dollar costs are indeed an important utility variable, but they do not necessarily dominate other variables. He is at least as concerned with the effectiveness of what the system does as he is with various dollar cost efficiencies. Dollar costs will be a constraining variable. The achievement of new system performance objectives may often tend to be dominant, assuming that they can be created within dollar constraints. The conceptual model of the executive making design decisions for changing the training system should be capable of a higher level of abstraction from existing system realities than is required, or even desirable, for the system control executive. The designer must be able to conceive of alternative possibilities to the present realities as potentially desirable, preferable, and even necessary. The designer's conceptual model should reference the major operational variables in the services to be provided by the system. The system controller's conceptual model is more likely to consist of the operating entities in the system and measures of their efficiency.

Reference models and procedural formats that are effective for making decisions should also be useful for learning to make a class of decisions. This instructional capability can be especially important for transitioning an executive from one type of conceptual reference for making decisions to another which demands a somewhat different conceptualization of the training system and a somewhat different set of values. Let's hope that it is not presumptuous to suggest that the reference content and the decision-making formats in ETAM can serve the purposes of orientation and of reorientation to decision making that deals with innovations to the training and using systems.

Notice that no assertion has been made that decision making can be "taught." This would invite unnecessary controversy. The assertion that decision making is learned, and that both its learning and performance can be helped with various tools and aids seems readily defensible. This argument should be a major justification for trying to make what is traditionally intuitive behavior at least partly explicit and its results communicable.

THE DECISION PROBLEM EXPRESSED IN LAYMAN'S TERMS

It can be illuminating to examine the decision problem in terms that make sense to the executive who actually makes decisions that involve risk and that will affect several variables in his operation at the same time, but who has been making them without benefit of formal utility theory.

Presume that background information has been developed about a proposed innovation for a training system. The executive may be offered three alternatives for which data have been developed:

    Accept the proposal
    Reject the proposal
    Accept the proposal tentatively by following a
        prescribed program of further study

The preliminary sample data about the innovation show that (1) it will make some reductions in operating costs; (2) it increases "student motivation to learn" in some observable functions such as increased class attendance, more time spent in study, and so on; and (3) increases instructor flexibility in kinds of subject matter that can be taught after brief orientations.

If the executive accepts the proposal, an investment is involved, and there can be three major types of outcome. If things work out well, the innovation may result in the savings and other benefits promised for it. If it fails in the process of being implemented, not only may costs accrue but operational liabilities may be incurred. If the project fails before it is implemented, only lost costs may be involved. If the executive rejects the proposal, the outcome is neutral--at least in the short term.

Now the study of the proposal may have led to the estimation of probabilities for each of the three outcomes if the executive accepted the proposal. This risk factor is, of course, one influence on the attractiveness of a decision to accept. Other things equal, the greater the estimated probability of success in achieving the promised benefits, the more appealing is the decision to accept the proposal.

The degree of risk will, of course, interact with the size of the commitment that is made by the executive if he accepts the proposal. The larger the investment, within the context of the executive's sphere of discretionary action, the less attractive is any level of success probability that is less than perfect.

Ideally, the executive should keep separate his picture of the risk associated with an outcome from his picture of the desirability of the outcome itself, if the outcome occurs. That is, the goodness or worth of the outcome should be evaluated in its own right as a set of conditions with some degree of

desirability. But we have three benefit variables if the proposal works out successfully. Let us say that, considered as a return on investment, the "reduction in operating costs" will not be so great as to justify an acceptance of the proposal. It must have additional benefits. The crunch question is that of determining the respective worth of these benefits so that they can be added to the worth of the moderate reduction in operating expenses. The objective is to obtain a figure of overall worth for this outcome.

Presumably the decision maker has information in his head about the values of benefit variables and incremental changes in those variables as they may affect his jurisdiction of operations. On reflection, he can project an image of the meaning, to his operation, of some incremental change in "student motivation to learn" from his knowledge of how his system is working with the present level of that motivation. (He should also be aware that this is a vague variable but that it connotes a set of observable events and conditions.) Perhaps the present levels of that variable are so low that the name of the variable itself is a "hot button."

A realistic examination of the actual magnitude of the incremental change promised by the innovation may help to moderate a precipitate commitment because the word "motivation" is connected with the benefit.

So the decision maker may think to himself: if the innovation works out, the increase in student motivation would be worth at least twice the value of the dollar savings promised by the innovation. The executive has now scaled the relative importance of these two benefits--where a "benefit" means the combination of a benefit variable and a magnitude of change in that variable.

He may have mixed feelings about the worth of the increase in instructor flexibility. He may think it has about the same worth as the very moderate dollar savings, but perhaps only one-fourth the worth of the increase in student motivation. He is now being inconsistent with his earlier judgment that the student motivation benefit was worth twice the dollar saving benefit. When this inconsistency is pointed out to him, he may reflect further and reach consistent answers.

Now the decision maker combines in his head the evaluations of magnitude and relative importance of the several benefits and implicitly assigns some conceptual figure of worth to that picture. Either now or later, he may figuratively multiply this composite worth picture by the risk picture and thus get a combined worth expectation for that outcome.

He may now examine the benefits and liabilities for each of the other two (or more) potential outcomes projected from a given choice point. He will follow the same procedure for estimating the worth of the benefit (or negative worth of a liability) in terms of incremental changes in the operational system. He will then also estimate the relative importance of the secondary benefits with respect to the primary or dominant variable and thus develop a more or less coherent picture of the composite of worth of this other outcome. The overall worth for the outcome may be conceptualized as the composite worth multiplied by the estimated probability of the outcome actually happening.

The executive's reference baseline for choosing an alternative course of action will almost certainly be the projected outcome of rejecting the

innovation: taking no action at all in the matter. This outcome may appear to him as a set of zero deviations on the benefit variables that appear in the innovation, and a zero investment commitment. Or the outcome from rejecting the decision to accept any proposal for action on the innovation may appear as progressive changes on those variables, in a plus or minus direction, over time on the basis of influences other than from the application of the innovation.

It is quite likely that many executives, when presented with a decision-making situation, do not abstract "variables" out of the problem and think of them as a continuous yardstick along which to measure magnitudes. This will seem to be a completely artificial operation. Rather, the executive who has operational knowledge of the domain he administers will think about the phenomenal changes promised by the innovation in the operational behaviors in the organization. These may be images of events, conditions, driving and control factors, dynamics and tensions among concrete operations in the training system.

We believe that this kind of concrete, phenomenal thinking should not be lost to the actual making of real world choices among alternatives. The ability to think realistically in terms of phenomena, actual or projected, may be central to effective intuitive decision making: the wholistic rather than the decomposition method. The latter is the analytic procedure for developing utilities and relative importance factors on individual and combined benefit variables and magnitudes.

A liability in any analytic method is that the scaling processes lead to progressively more abstraction and thus remoteness from the real world of complex manifestations and dependencies, lash and backlash, and overall balance in the behavior of complex systems such as the training organization and its interfaces. The abstractions, such as the utility function for a complex variable like "motivation to learn", may become reified--treated as if they were indeed real yardsticks measuring magnitudes that could be converted into useful judgments of the worth of the incremental magnitude change. Thus, while the analytic procedure may gain information in the form of the assessor's and executive's scales of worth and importance of the dimensions and attributes of the system being administered, the analytic activity also loses information about the real world of process and process dynamics to which the consequence of the decision will apply.

Ideally, of course, the wholistic and intuitional route and the analytic, decompositional route would lead to the same conclusions. Practically, the decision maker can be overloaded with too much information so that his thinking becomes (1) stereotyped, (2) random, or (3) limited to a subset of the relevant decision data. Hence the desirability for analytic decomposition of the problem into components. But fiddling with the components may be like looking at the trees rather than the structure of the forest when planning a firefighting campaign. Therefore, intuitional and "common sense" judgments should be interspersed with analytic and synthetic operations with the explicit strategy of convergence.

This strategy tends to bypass a variety of technical difficulties that beset the decompositional procedures. Fisher (1974) identifies some twenty-eight procedural models of varying complexity. Any decision model procedure depends on some degree of conformance to various assumptions: additivity, linearity, independence, consistency and so on. The determination of whether one or all

of such assumptions are sufficiently valid to justify a given procedure calls for time-consuming and often quite artificial operations. The greater the number of these tests and constraints imposed on the assessor or decision maker, the further he gets away from the operational guts of the real problem he is trying to solve.

We should also recognize that in this project our concern is not with research into the models that can account for the decisions that decision makers make, nor a concern with determining what ultimate level of precision in characterizing choices can be derived from any given content of information either inside or outside the human head. Our objective is to provide a procedure that generates (1) the characterization of decision pathways and outcomes with precision no more than sufficient to identify to the choice maker the distinctions in significance between one projected outcome and another, (2) a useful level of specifying the risks associated with a given choice, and (3) a practical basis for ironing out disagreements among various participants in a decision-making process.

In any event, the wholistic and intuitional methods can always give the perplexed user an escape hatch in the event that no decision model seems quite applicable to the problem, either in terms of assumptions or convenience, or because of the complexity and importance of the problem to be solved.

This strategy of interleaving wholistic and decompositional activities in formalizing a decision is followed in the procedures recommended in Section II.

PART B. RATIONALES FOR THE DECOMPOSITION OF COMPLEX DECISIONS

In this section the major liabilities and limitations of wholistic and intui-
tive decisions making will be summarized and related to assumptions and operations
in assessment and scaling procedures. These psychological limitations are
presented with the assumption that utility and probability scaling techniques
can counteract some of their influence. The major supporting references for
the following content are Guilford (1954), Kahneman and Tversky (1973), Baumol
(1965), Goldberg (1968), Spetzler and Stael von Holstein (1972) and Huber (1974).

INFORMATION OVERLOAD

This is a general catch-all characterization of the inability of the human to
hold in mind and process more than a limited amount of information. If the
decision has several alternatives, and each alternative has several potential
outcomes, and each potential outcome has a number of benefits, liabilities, and
probabilities, the amount of information swamps the decision maker's ability
to cope. George Miller's (1956) article asserting that the human can handle no
more than five to seven chunks of information at a time is cited, sometimes
inapplicably, as a limiting condition.

Slovic and Lichtenstein (1971) say that the number of stimulus attributes which
people can systematically consider is from five to ten. Huber (1974) states:
"Although we have successfully had clients respond to eight-attribute items,
they were so taxed by the task that it is doubtful that we could have gotten
them to perform it again." Huber is unclear about the procedural context in
which the task was performed. He does suggest careful "panning" of decision
variables but acknowledges that procedures "for so doing" have not been tested
for their strengths and weaknesses.

The overloaded human may follow any one of several courses. One is to over-
simplify the problem by attending only to those factors with which he is most
familiar, or appear most significant to him. Other factors, including qualifi-
cations of those he is considering, are neglected.

Because constructive thinking is facilitated by concrete ideas, tangible factors
may predominate over more abstract factors when the human is overloaded. Because
dollar costs have "tangible" references, they may consistently dominate the
overloaded decision maker.

In information overload, the simplifying process may seek a postponement of a
commitment by deferring the actual decision beyond the point where it is profit-
able to make it, even when there is sufficient information to make a choice.

The assessment of risks will tend to be non-rational. The confused person may
make an arbitrary choice by the equivalent of tossing a coin; or because of his
uncertainties, he may take an unduly conservative position which is that of
rejecting proposals as a defensive strategy.

The overloaded person will improperly evaluate any additional data offered as

evidence pro or con for a decision-making situation. He may overvalue the weight of objectively trivial evidence, and fail to distinguish between relevant and worthless evidence. Thus he may be overly responsive to any apparent change in his perception of the problem situation (Goldberg, 1968).

INCONSISTENCIES IN EVALUATION BECAUSE OF ABSTRACTNESS OF VALUES

A wholistic assessment of a multi-attribute decision runs the risk of splitting off the perception of expressed size of a benefit from the perception of the operational meaning of the benefit. This occurs because the intuitive decision maker is having to combine a number of benefits into a single conglomerate of worth. In an earlier section of this report we have seen the several conceptual elements that translate an incremental change in a benefit variable into a "system goodness utility." This utility continuum ranges below and above the present state of the attribute with a floor consisting of a Least Acceptable State and a ceiling consisting of an Ideal (or Goal) State.

By forcing the decision maker to convert one benefit at a time into a position on its utility scale, the decision maker can attend to the contextual meaning of the benefit attribute and the benefit magnitude with respect to its distance between a present state and either the Least Acceptable State or Ideal State of the attribute. The development of the incremental value of the benefit from the description of the benefit reduces the vagueness (and unreliability) of what otherwise may be unanchored abstractness.

This theme will be repeated when we examine distortions of values and inconsistencies in judgments among values.

DISTORTIONS IN PERCEPTION OF VALUES

The decisions of the wholistic decision maker are subject to a variety of distorting influences. Recent dramatic occasions may unduly sensitize him to a given system variable so that any decision in which that variable appears may become dominated by the remembrance of the dramatic occasion. This distortion may extend to an entire class of decision contexts so that he is predisposed to an Accept (or Reject) decision for any proposal which contains the key word or concept.

Another liability is the halo effect. This is the impression that if an entity has shown some one highly desirable attribute, all the rest of its attributes will also be desirable. Or, conversely, if the entity has exhibited some undesirable traits, its other attributes may be perceived as undesirable. This atmosphere effect can pervade the entire context of a decision problem and predispose the decision maker to overall attractiveness or unattractiveness of the proposal even before a proper examination has been made of it.

Decomposition techniques tend to offset these tendencies in two major ways. One is by forcing the decision maker to examine and evaluate each benefit or liability apart from the evaluation of other benefits so that the dominance of an atmosphere effect is at least diluted. Each attribute has a value dimension in its own light and according to the evidence for its magnitude of change. The second advantage is that by examining and evaluating attributes one at a time, the decision maker is not so overloaded that he cannot also attend to the data summaries that provide substantive evidence of the individual benefit. To the extent that each attribute for each projected outcome can

vary independent of the influence of other benefits, the overall variations between projected outcomes of decisions become larger. This is certainly desirable from the standpoint of emphasizing the difference in overall worth between one outcome and another, and thus giving appropriate confidence to the choice of a given alternative.

INCONSISTENCY OF EVALUATIONS

There are many ways in which the decision maker can be inconsistent without its being detected in wholistic decisions. Baumol (1965) identifies a number of assumptions about the "rational" decision maker that may readily be violated. Thus: transitivity--if an individual is indifferent to prizes X and Y, and he is indifferent to Y and Z, he should be indifferent to X and Z. Continuity of preference--if he prefer X to Y and Y to Z, he should prefer X to Z. Independence--if the individual is indifferent to getting some incremental benefit X or Y, he should be indifferent to outcomes that give him either an X or a Y. Desire for high probability of success--given lotteries with equal prizes, the individual will prefer that with the highest probability of winning.

Decomposition methods more or less force consistency on the decision maker by requiring him to sequence choices and preferences in an order that either prevents inconsistencies from occurring, or makes them manifest when they do. Note the style in the procedures which require the scaling of utilities and of relative importances for each benefit and liability. Common reference bases, either using a continuous base variable such as dollars, or an arbitrary scale such as system utility, enable assessor to apply a consistent point of view to diverse decision attributes.

WARNING

It is indeed quite possible for benefits to be ranked in the order of X, Y and Z in one context and legitimately ranked in Z, X and Y in another context. Because of this possibility, it is desirable to have the contextual facts about the benefit available for inspection along with the mere naming and quantification of the benefit. This context is given by the range of applicability of the innovation.

Another basis for inconsistency can be a shifting picture of what variables and functions have higher and lower priority for improvements in the training/ operational system for which innovations are to be considered. To the extent that these priorities are consistent and firmly established, judgments that reference them will be consistent.

But because the prioritizing of benefit variables for relative importance may be difficult--because it deals with abstractions--judgments about relative importance in ETAM procedures compare one level or incremental magnitude of benefit variable X with another incremental magnitude of variable X (in scaling the relative utility of different values of benefit variable χ) or with variable Y the relative importance in utility of the benefit--as a level on the benefit variable--within the same outcome package. This is like the difference between determining whether flour is more important than sugar, and asking: "Is getting three pounds of flour right now more important than getting four pounds of sugar?" The latter is likely to be a more meaningful

and answerable question because it references concrete contexts.

Another basis for inconsistency can be a shifting picture of what variables and functions have higher and lower priority for improvements in the training/operational system for which innovations are to be considered. To the extent that these priorities are consistent with each other and firmly established, judgments that reference them can be consistent.

But even under such conditions, the weighting of variables as such is difficult to do. One might ask: in the kitchen preparation of food, which is consistently more "important", flour or sugar? It would be easier for the cook in the kitchen to answer the following question: On the basis of your present supply of sugar and flour, and the meals you plan for the next few days, which would be more important right now, an additional two pounds of flour or an additional three pounds of sugar?" It is this latter format that is used in the scaling procedures we recommend for ETAM.

Still another basis for inconsistency in global judgments is that in the case of some attributes, the present or reference state may be at the very high end of the range or at the very low end of the range. In such cases, a given increment of change may have much greater significance than the same increment of change from a middle level present condition. In the decomposition treatment, the utility value of incremental change on each attribute is determined before it is aggregated with other utilities. See Steps 5-6, and 7-8 in the procedures, Section II, Figure II-3 of this report.

UNRELIABILITY IN THE EVALUATION OF QUALITATIVE ATTRIBUTES

The subjective translation of some incremental change in a qualitative attribute into an expression of desirability (utility) is notoriously vague and hence unreliable. The unreliability stems from two major factors. One is an unclear operational reference for the qualitative attribute; the other is an unclear or shifting value reference from which to select a measure of desirability. In wholistic thinking this difficulty is compounded when the qualitative attribute (for example, level of "professionalism" among instructors) is one among a number of other benefit attributes promised by an innovation.

In ETAM procedures vagueness in the operational reference is reduced by requiring the assessor or decision-making surrogate to specify at least a sampling of the kinds of observable operation denoted by the name of the attribute and applicable to the innovation. Thus, a reference list of operations associated with "instructor professionalism" might include: spends extra time in studying the material to be taught; goes into the field to get realistic examples; adapts instruction content to needs and capabilities of individual students--and so forth. The incremental change promised by the innovation on this attribute in one projected outcome may have been summarized by the scalar statement: "moderate increase in instructor professionalism." The incremental change projected on this attribute in another projected outcome might be "small but noticeable decrease in instructor professionalism."

The second factor in the unreliability of wholistic judgments is in a scalar reference for converting several incremental magnitudes of the attribute into differences in a scale of desirability. In this case, how much is it worth to gain the moderate increase in instructor professionalism as contrasted with

the worth of avoiding a small but noticeable decrease in instructor motivation. Let's assume that these two outcomes are the best and the worst among the set of outcomes projected for the decision alternatives and action pathways. The principle of tradeoffs can be applied for determining relative worth.

The procedure for making this translation should be the simplest possible because it is difficult for the assessor to hold both contexts in mind while engaged in the judgmental process. The rating or variable interval demands of the lottery method would not be indicated. The lottery method requires the judge to identify the odds placed on each item which he would accept in a bet. The ratio of the odds is the ratio of value between item 1 and item 2. The latter method does require the person to examine his preferences in a way not required by the rating method.

The judge will also be helped by referring the incremental change represented by each benefit to the range limits on the benefit variable: the Least Tolerable level at one end and the Practical Ideal at the other.

CONFUSING THE GOODNESS OF A BENEFIT WITH ITS PROBABILITY OF HAPPENING

The classical example of this confusion is expressed by the fox who, sizing up the improbability of his leaping high enough to get the grapes, concluded that they were sour anyway. This tendency for a subjective estimate of a probability to influence the estimate of worth of some goal or objective can distort the rational evaluation of decision outcomes, and thereby impair the quality of the decision.

When the scaling of the utility or desirability of a projected outcome is separated from the process of determining the subjective probability of achieving the outcome, this tendency to confuse one with the other is reduced. This separation of utility scaling and probability estimation is maintained in the ETAM procedures: in Figure II-3 Steps 1 through 9 deal exclusively with the determination of utilities and their summation into utility values for projected outcomes. Probability estimates are developed in different contexts in ETAM Tasks, 3, 4 and 6 and are combined with outcome utilities in Step 10.

The overall worth or expected utility for a given outcome is its subjectively estimated probability multipled by the estimated utility of the outcome--where utility is a scale of something desired: equivalent dollars, satisfaction, system goodness. By keeping the utility and probability factors separate, it is possible to make changes in one of these factors, hold the other constant, and compute the changes in the expected utility of the respective decision alternatives. If probability estimates are changed, it is possible to determine how sensitive the outcomes (and the attractiveness of given alternatives leading to those outcomes) are to probability estimates. Conversely, changes may be introduced into one or more of the benefit attribute levels, and be translated into corresponding utility values. The overall worth of the projected outcomes is recomputed, and the sensitivity of the decision to any given benefit attribute is determined. As a consequence, the directions are specified for the most fruitful further inquiry preceding a final choice of alternatives.

These kinds of manipulation with the information comprising the information content of a decision are virtually impossible with an exclusively wholistic treatment.

A warning may be in order with respect to the separation of the utility of an outcome and the estimated probability of achieving the outcome. The lottery process of scaling the values of incremental magnitudes of a benefit attribute may conceivably lead to the subject's confusion of the probabilities used for scaling utilities with the probabilities used for assessing the expected utility of alternatives. It may therefore be important that training sessions correctly orient the person on the difference in contexts.

LACK OF JUDGMENTAL VARIATION IN PREDICTING VAROUS OUTCOMES

Wholistic judgments in assessing multiple-attribute outcomes projected from decision alternatives may tend to fall consistently into the middle ranges of goodness. It is like a number of colors superimposed that produce various kinds of gray. The potentially real differences between outcomes may be conceptually blurred so that overall judgments are regressive towards neutrality such as "outcome A looks moderately good and outcome B which admittedly differs from A looks moderately good, too." This generalization will have its exceptions.

The decomposition method and scaling techniques elicit informa' n from the decision maker which tend to help the perception of real diffei ⌐ces in worth among various outcomes if the analytic judgments of the decision maker warrant this emphasis. By scaling each benefit attribute and its projected increment of change into a utility index that has a zero reference ("the least acceptable level of the attribute"), there is reasonable justification for adding or subtracting the several utilities in a multi-attribute outcome. (It is "reasonable" if the linearity assumption is reasonable here.) Thus, the value of incremental changes are added into a single dimensional yardstick. At least in theory, this cumulative expression of merit on a single scale should enhance whatever judgmental differences exist between the overall goodness of outcome A contrasted with outcome B.

The foregoing argument does have some holes in it. The knowledgeable and sensitive decision maker may be able to distinguigh major differences in the implications of overall patterns of change introduced into a system as contrasted with another pattern of change. But analytic techniques on the respective components of each pattern may fail to show any important differences among them. For example, on a feature-by-feature examination, two faces may appear to be quite similar to each other, but seeing the entire faces as patterns makes them extremely easy to differentiate. This analogy suggests, however, that a combination of both methods would be desirable for the enhancement of the perception and recall of differences, as well as for the recall of similarities between the faces.

Decision making is simplified and achieved with greater confidence when the alternatives lead to large differences in the utility of different outcomes. Any legitimate basis for enhancing the differences among projected outcomes is therefore an aid to the reliability and probably to the validity of the decision process. (What is similar or common to all outcomes of all alternatives should cancel out of the problem and not clutter it further.) The foregoing arguments suggest that a combination of the wholistic and the decomposition methods may be desirable for the emphasis of legitimate differences among projected outcomes. The analytic process should sharpen the perception of the contribution of component outcome attributes singly and laid end-to-end. The wholistic examination of the pattern of attribute change projected for a given

outcome may lead to the perception of an important balance of factors.

This last comment suggests that in some problems and the operational contexts which they reference, variables are not independent of each other. This is a serious topic in its own right, and is treated next.

PERCEPTION OF INTERACTIONS AMONG ATTRIBUTE VARIABLES IN THE SYSTEM

In general, the simplicity and practicality of utility scaling assumes that incremental changes in one benefit have increases in goodness that are indepen-dent of the incremental changes in other benefits. This assumption is important if units of goodness derived from incremental change on one attribute are to be added to units of goodness of incremental change derived from another variable. It is the challenge to this assumption of independence of attributes that is perhaps the major objection of the wholistic and intuitive decision maker to analytic techniques. Most of us believe that we respond to configu-ration of data rather than to a set of data items.

Unfortunately the evidence for the value of configural judgments among variables in decision making by humans is poor. Goldberg (1968) cites the results of studies where the human subject believed he was using interactive relationships for problem assessment. But the results of the subjects were no better than could be accounted for by additive (analytic) models that assumed independence of the attributes. Since studies of this kind are limited to available samples of persons and problem situations, the reader might choose to hold out for more extensive kinds of evidence before abandoning his belief that he responds to the more or less unique pattern of attributes and attribute values in assessing an outcome. The researchers do not assert that humans don't make configural judgments, but only that they don't have other results than are obtained by simpler additive models.

We recommend playing it safe. Notice that in the procedures of Section II, Figure II-3, Steps 5-6 and 7-8, the analytic method is used. But also notice that Step 9 calls for consistency checks that include global judgments about the relative goodness of the combination or pattern of benefit attributes and their incremental change levels between one projected outcome and other projected outcomes. This is essentially a wholistic operation. Step 9 calls for recycl-ing of the analytic and wholistic operations if they have conflicting results with respect to the evaluation of the different projected outcomes.

The value of the interactive computer in these operations is in enabling the user quickly to compute a set of results, and where two sets of results do not match, quickly to identify where in the procedural trail the mismatch was generated. With a computer, it is simple to make tentative or trial changes and recalculate results before making commitments. The labor in equivalent manual operations would make these kinds of convergent adjustments impractical.

LIMITATIONS FROM BEING RESTRICTED TO ONE POINT OF VIEW

The complexity of wholistic decision making generally tends to limit who will or can undertake to do it. If more than one individual makes assessments of the same problem and they come to radically different evaluations of projected outcomes from decision alternatives, there is little basis for their converging

towards agreements if they all need wholistic approaches.

Goldberg (1968) and Chapman and Chapman (1967) cite data showing the power
of pre-established biases and pre-judgments on the assessment of evidence
and conclusions drawn from the evidence as they would apply to clinical
decisions. At least tentatively, such findings may be extrapolated to other
classes of decisions.

Some studies, such as dealing with weather predictions, have shown that where
procedures foster their working together, multiple heads can often be better
than one in assessment and interpretations leading to decisions. (Murphy and
Winkler, 1973.) Whereas some studies show the maximum increase in goodness of
prediction comes from two rather than one person, with rapidly diminishing
returns from adding more people, other studies imply that the goodness of
consensus increased appreciable up to eight participants. Probably the greater
the expertise of each participant, the fewer participants leading to markedly
diminished returns.

Many more opportunities exist for constructive interaction among several partici-
pants in all stages of the assessment process in the decomposed than the wholistic
method. Furthermore, if an specific disagreement is reconciled when the decom-
posed method is used, the change is easily entered and the calculations readily
recomputed.

Two or more assessors may combine their efforts both in the determining of the
utilities which combine to establish the value of given projected action out-
comes, and in the estimation of the subjective probabilities of their occurrence.
The participants may begin by working independently through the procedural
Steps 4 through 9 before comparing their final conclusions and the data developed
in each component step and substep in the analytic procedures. If they begin
by working independently, their starting differences may be one diagnostic sign
of the uncertainties in the problem and the data generated in the problem. For
some purposes, it would therefore be desirable to maintain a historical record
of the individual efforts, and the rationales that are used for convergence
when consensus is reached, or for divergence if consensus is not reached. This
record could influence the decision maker's subjective certainties about the
validity and precision of the assessments presented to him.

It would appear highly probable, therefore that there would be a greater tendency
for assessors to agree in their final judgment if they all had equal access to
the same specific, concrete facts throughout the evaluation.

Stael von Holstein (1974) has developed and formalized procedures for obtaining
probability estimates and distributions from groups. The procedures include
operations for achieving consensus.

BIAS TOWARDS PESSIMISM OR OPTIMISM

Our subjective experiences with our associates lead us to characterize some
of them as optimistic, others as pessimistic. In some cases, one or the other
attitude appears to dominate the individual's assessments, decisions and
interpretations of his experiences. The attitude may be limited to the role
an individual may be playing at a given time, or it may be a cross-sectional

personality characteristic. In the absence of other information, we should
perhaps assume that levels of optimism-pessimism are distributed normally in
the population.

The effects of a consistent optimism-pessimism bias can be pervasive and
highly influential in assessments and decisions. Like other forms of bias,
its influence is exerted primarily where there is uncertainty and ambiguity
in a problem situation. There is always uncertainty and ambiguity in assessing
and deciding about innovations. An optimism-pessimism bias can be compounded
if the assessor who interprets and casts up the results, and the decision maker
who interprets the presentation, both have a bias in the same direction. At
least the assessment process should be neutral as possible.

Optimism-pessimism bias can show itself in many ways. A pessimist can set
the mean value of a projected benefit from an innovation too low. This is
possible because the projection is inevitably based on sample data from sample
conditions that must be extrapolated to real life conditions. The pessimist
can skew his confidence range within which the mean value of the benefit is
expected to fall in that he projects a larger reduction of the benefit to
happen under "adverse conditions" than he projects an increase in the benefit
under conditions "highly favorable" for the benefit. The pessimist can put
greater weights on liabilities (negative benefits) than on positive benefits
in terms of their magnitudes, ranges of effect, and their utility or negative
worth, and on their probability of occurring (Slovic, 1966).

The optimisitically biased assesssor will tend to make evaluations that have a
pattern which is the reverse of that described for the pessimist.

The decision maker expresses optimistic-pessimistic bias in the form of being
risk-accepting, in which case he follows a policy of maximizing his potential
gains, or of being risk-averse, in which case his policy is that of minimizing
his potential losses.

One way to reduce unwanted bias is to get consensus where the judgments of
individuals with bias in one direction are probably offset by those with an
opposite bias. One can only hope that this correction will happen with
the sample chosen.

The effect may be reduced by simplifying the assessment and decision-making
procedures, and keeping judgments as close to reference facts as possible all
along the line. This argues for analytic and reductionist procedures, and
quite definitely against wholistic procedures which heavily depend on "intuitive"
global evaluations where systematic bias can most readily operate.

A second method of coping with optimistic-pessimistic bias could consist of an
independent measure of this tendency, such as by self-assessment on this trait,
or assessment by others, or by responses to a specially designed and validated
test questionnaire. Where a pervasive bias is determined, and it can be
measured, the results of judgments in the assessment-decision context can have
"correction factors" applied to them. The use or development of specialized
tests is, however, outside the scope of this investigation.

Hurwicz has investigated this topic and developed an optimism-pessimism "index"
that can be applied to projected outcomes in decision situations. The method

is criticized by Fishburn (1966). Further methodological inquiry seems indicated.

In any event, the _assessment_ of information content that enters into what is presented to the decision maker should be as neutral as possible with respect to optimism-pessimism bias. If such a bias is to be applied, it should clearly be the prerogative of the decision maker himself to exercise it in the selection of an alternative course of action.

MISPERCEPTION OF THE TIME HORIZON OF THE PROBLEM

The development, testing and introduction of many innovations may take months and years. The operational realization of benefits may begin small, increase to some maximum and then decline. So the assessor and decision maker must depend on some combination of facts and intuition as to the worth of the projected benefit over a time span that is perhaps equivalent to the amortization of the capital investment on one hand, and on the other hand to the projected needs of the training and operational systems.

The time scale that applies to the decision should also take into account the progressive changes in present conditions if the benefits of the innovation are not introduced.

Adding the time dimension to the many other dimensions of the assessment-decision task puts another burden on thought and judgment. This burden may be substantially reduced by the analytic examination of each benefit variable apart from the others in determining both the projected level of the variable _without_ the innovation and the projected level _with_ the benefit across the time span being considered.

By taking each benefit or liability into account separately and evaluating its specific contribution to system goals, different time horizons can be appropriately applied.

DETERMINING RELATIVE IMPORTANCES OF BENEFITS/LIABILITIES FOR A SINGLE FIGURE OF MERIT

It is in this step--where several benefits, some of which may be negative, are given different weights and combined into a single figure of merit for an outcome--that all improper assumptions and procedural difficulties become magnified. All scaling issues come to the big crunch in this step. If a weighted magnitude on variable X is to be added to a weighted magnitude in variable Y and be translated into some continuum variable such as general worth, consistent biases in calculating the weights or in scaling the variables will summate into heavily biased totals. If the reference continuum is inappropriate to the requirements of the decision context, the outcomes presented may be irrelevant to the decision that ought to be made. If the interpretation of the scale values, such as equivalent dollar utility, or satisfaction, or system goodness utility, assumes a zero point which does not exist as a reference, or which is misplaced, the presentation of the various outcomes to decision alternatives will be misleading. Where ther are many perils, wholistic common sense should certainly interact with analytic techniques in producing the final judgment.

Because of its complexity, it is this operation which is most likely to overload the decision maker if there are more than two benefit variables. And human processing liabilities in overloading will occur: preferential bias for a given variable, halo effect, randomness in referencing, selective sampling of relevant information contained in the problem, primacy and recency of what was considered, and so on. Thus it is in this step that explicit decomposition procedures can be of greatest help to the decision maker.

There are several kinds of relative importance that apply.

a. The relative importance of different incremental magnitudes of the the same benefit variable for different projected outcomes.

b. The relative importance of each of the different benefits associated with any one outcome.

c. The relative importance of outcomes to each other.

The determination of relative importance of benefits and the determination of utilities for individual and combined benefits is an interactive procedure. In principle, the assessor who has examined the operational significance of the various benefits and liabilities for each projected outcome makes an intuitive conversion of each benefit or liability into some common utility (dollars, system goodness, satisfactoriness, or whatever). He then selects the "most important" benefit according to this intuitive assessment. There is also a practical criterion for this selection. The incremental range of the "dominant" benefit when translated into a utility scale (formally or intuitively) should be greater than the incremental range of any of the other benefits. That is, its range should embrace the utility ranges of the other benefit variables in the "benefit package."

Ideally, the range of this reference benefit variable should be somewhat larger than sufficient to encompass the ranges of the incremental changes in the other benefit variables across all projected outcomes. If the range of the reference or dominant benefit variable across the values it assumes in the various projected outcomes is very much larger than the next most important variable (say, more than two or three times) then it may cease to be a reliable guide against which to scale the other benefits. Within limits, the more similar the magnitudes to be compared with each other, the more reliable the comparisons, and the greater the dispersions of the resulting judgments. For example, objects that differ from each other by few square inches are more readily scaled in terms of their proportional differences when a foot square is used as reference than when a yard square reference is used.

Other things equal, the more tangible the reference variable against which to weight the importance of the other benefits (as variables and as incremental magnitudes) the easier and more reliable the judgments should be. It is perhaps this tangibility factor that might tend to favor a costable variable--such as dollars--as a primary reference variable. Furthermore, in the decision contexts we are talking about here, it is somewhat more likely that there is a straight line relationship between dollars and utility for dollars within the ranges likely to be influenced by a given innovation. This could, however, be a

dangerous assumption to make, and it should be accepted very tentatively.

In summary, the following are the major alternatives in procedure for determining the relative importance of multiple benefits and multiple outcomes.

a. Develop the intuitive expression of a wholistic image of worth for each whole outcome in some symbolic terms that may be quantitative or qualitative.

b. Use a dominant costable benefit variable as a reference against which to weight other benefits at each outcome, and as a metric into which to convert those weights and benefits as a figure of merit for that outcome and other outcomes.

c. Use some other continuous benefit variable that is dominant as a reference for developing weights and conversion of individual benefit magnitudes into a common utility scale such as "satisfactoriness."

d. Use a "system goodness utility" reference for developing weights of individual benefits and for converting into a common utility scale.

The advantages and liabilities of the wholistic treatment have been discussed.

The selection of a costable variable slants the entire evaluation of outcomes towards a monetary reference with a consequent risk of slighting the values of operational system goodness. Nevertheless, this can be for many decision makers and decision contexts a meaningful orientation for utility. One justification is that many decision makers have traditionally used this reference, and are experienced in thinking with it. If the decision is important both in terms of financial commitment and in terms of changing mission objectives, two independent but parallel evaluations would seem justified: one with a dollar utility reference, the other with a system utility reference. If the results developed from each of the two scaling references can not readily be converged, it should be clear that policy reconciliations are necessary to clarify the organization's scheme of worth that relates dollar utilities to system utilities. Schlaifer (1967, 2-13 to 2-22), argues cogently for the conversion of nonmonetary consequenes to monetary criteria in business situations.

A continuous benefit or liability variable that is dominant in the analytic evaluation of the innovation may be used as an alternative either to a dollar cost or system goodness reference for developing relative weights. Thus, one would expect that an innovation directed primarily at increasing "instructor competence" would have this as its dominant variable. (Recall that the assessor will have provided samples of operational meaning for this abstract variable.) The contributions to each projected outcome by other variables (such as for example "student interest in the subject matter") can be weighted against the non-costable variable for relative importance. Such a non-costable variable, other than system goodness, may be useful for an organization which provides ancillary services to the overall training and operational system, where it would be difficult to apply the system goodness

reference in a meaningful way. The utility reference for the non-costable dominate variable might, in this example, be "satisfactoriness level in student-instructor effectiveness."

The system goodness reference would seem ideal, but like many ideals, it may often be impractical either because of its vagueness to many decision-making entities, or because the projected effects of the innovation have important local applicability that is not of sufficient scope to apply to the system goodness range. It would certainly be desirable to develop some empirical information about the practical usefulness of the concept of system goodness as a reference utility measure for decisions about innovations that had targets other than displaceable costs.

Eckenrode (1965) has experimentally evaluated a number of commonly used procedures for weighting multiple criteria. These procedures consisted of ranking, rating, partial paired comparisons, complete paired comparisons and successive comparisons. Most of the intercorrelations among the procedures were extremely high: in the middle and high nineties. But the rating and ranking methods required only a half to one-sixth the time to perform as the other procedures, and were preferred by the subjects. Rating was accomplished by having the subject draw a line on a value scale ranging from zero to ten, to any point appropriate for the variable under consideration. We have recommended the rating method with appropriate checks for consistency, as the primary basis for judging the relative weights of benefits at each projected outcome. Huber (1974) also favors this method as being easy for his clients to use.

Some researchers and practitioners have used a lottery method for scaling relative importance of differenct values of the same variable or of given values of different variables. In principle, the judge is presented with different hypothetical probabilities (P) of getting benefit X and 1-P of getting benefit Y (or avoiding negative benefit Z) until acceptable odds are presented. If he would take a gamble on an 80% probability of getting X and a 20% probability of getting Y, the importance of X is four times the importance of Y.

An alternative format for the lottery method is to get the judge to select odds that he would accept for getting X as contrasted with getting Y. For example, the judge might accept odds of three to two in favor of getting X. This gamble would be translated as a value of 3/5 for X and 2/5 for Y. This format seems more difficult for many judges because it requires more mental processing. One should be reminded that the greater the mental difficulty in performing a procedural operation, the less reliable its outcome may be (other things equal) and the greater the chances that the judge loses sight of the task he is really trying to get done in the mechanics of procedural complexities.

The lottery method for establishing relative importance among elements serves only to elicit judgments of relative weights and preferences. The operation has nothing to do with the determination of subjective probabilities of outcomes. The psychological value of the method seems to lie in its forcing the judge to think seriously about the relative significance to him (or to whatever reference he is using) of getting X versus not getting Y, and of quantifying this significance. The selection of ratings and rankings

by the lottery method seems to invite more serious reflection about the judgments he is making. We have not found any research which correlates behavioral data obtained by this method with the other psychophysical procedures.

To the reader who has not experienced these several methods for weighting multiple attributes and attribute values, the foregoing procedures may seem complicated and time-consuming. In fact, however, they are not. Even a five or six attribute problem having four alternative outcomes to evaluate may be put through any two of these procedures in 10 or 15 minutes (or less), assuming of course that the judge does have well in mind the subjective impor- tance of the various benefits and the meanings of these benefits in an operational sense. These time periods are our own estimates based on some informal paper and pencil exercises.

A utility and decision-making support system should therefore enable an assessor to choose from among several scaling formats those with which he feels most comfortable (Huber, 1974). Perhaps he should be encouraged to apply more than one procedure for the same data in order to determine the consistency of his expression of judgments. The simplest procedures seem to be interactive graphics that would enable the judge to represent single dimension relationships by respective line lengths which he might draw or select on a screen. Two dimensionsal relationships, such as (a) the utility function between the incremental magnitudes of a variable on an abscissa and (b) utilities on the ordinate, might also be drawn on the screen. But such actions by the judge should be accompanied by suitable interrogations and cross-checks imposed by the information system.

A suitably programmed interactive system could find its most useful purpose in training the assessor and decision maker in the formulation and expression of the problem to be solved, and in useful ways to think about it while scaling the data into judgments of relative importance, formulating a single figure of merit for projected outcomes, and comparing outcomes. The training should be what Huber (1974) calls "client-centered:, that is, it should be directed towards the task of the decision team, and not to the mathematical logic underlying the techniques.

Such training should warn of pitfalls such as being preoccupied with method rather than the objectives of the procedure, and the importance of cross- checking complex judgments by using alternative procedures for processing the same data. One alternative should always be a common sense examination — at the wholistic level.

The most important single warning that can be made about all of these formal decision and utility procedures is that, in common with other techniques, the user may become lulled into believing that there is magic in them, and thus become an unthinking slave to them. This liability may apply especially to the user of material generated by these procedures: not only may he cease to credit his common sense and intuitions with any great value, but actually stop using them altogether. All of us have seen examples where the format tyrannized the operation.

This danger will be compounded by computer processing. To many people, including those that consider themselves enlightened, a computer output has an

atmosphere of inherent rightness and authority. This atmosphere is even stronger than that of the printed page or TV screen.

It would be wise, as well as ethical, for the training and procedural programs that are used in any aspect of utility and decision procedures to embody both the spirit and practice of these warnings.

THE ASSESSMENT OF PROBABILITIES

The preceding rationales were based on our recommended ETAM procedures which separate the descriptions of projected benefits for various outcomes from the estimation of the probabilities that a given outcome will happen, given the choice of a decision alternative. Practically every writer in utility-decision theory points out the logical dangers of mixing up estimates of probability of outcomes with the goodness or utility of the outcome during the process of evaluating either probability or utility. Slovic (1966) summarizes ways in which probability and value interact for some people.

To repeat, the previous sections have dealt only with utility scaling of outcomes which permitted the form of analysis applied to so-called "riskless decision." Thus, our procedure separates utility assessment from probability assessment. In ETAM procedures Task 3 and Task 4, operations are performed that lead to the assignment of probabilities for the several projected outcomes in a decision tree. It is in the context of judgments made in that task that the following brief remarks will be made. The literature on assessing probabilities, by the way, is far richer and more prescriptive than the literature on multi-attribute scaling in practical situations. Spetzler and Stael von Holstein (1972) at Stanford Research Institute have developed computer programs for conversational development of probability assessments. A comprehensive paper by these two researchers describes types of biases in probability encoding. The following is a summary of what they identify as sources of judgmental bias. A bias is defined by them as "conscious or unsconscious discrepancies between the subject's responses and an accurate description of his underlying knowledge." The forms of bias are:

Displacement bias: "a shift of the whole distribution upward or downward relative to the basic judgment."

Variability bias: "a change in the shape of the distribution compared with the underlying judgment--the distribution has less . . . spread than is justified by the subject's actual state of information."

Motivational bias: conscious or subconscious adjustments motivated by "perceived system of rewards."

Cognitive bias: systematic adjustments to the subject's responses introduced by the way "the subject is intellectually processing his perceptions."

Spetzler and Stael von Holstein go on to describe basic modes of judgment, and we continue to quote from them.

Availability: "Probability assignments are based on information that the subject recalls or visualizes."

Adjustment and anchoring: "The most readily available piece of information often forms an initial basis for formulating responses; subsequent responses then represent adjustments from this basis." Thus, recent experiences may dominate as a probability anchor.

Representativeness: " . . . the probability of an event or a sample is evaluated by the degree to which it is representative of, or similar to, major characteristics of the process or population from which it originated." Thus, people may erroneously assign the same distribution range to a set of averages that they assign to the distribution of events within a single average.

Unstated assumptions: frequently the subject's estimates at the start are based on limited samples in his own experience, and as he is guided to reflect further on facts and processes, he takes more information into account and may revise his probability picture.

Coherence: "People sometimes assign probabilities to an event based on the ease with which they can fabricate a plausible scenario that would lead to the occurrence of the event." The judged probability therefore depends on whether the subject can or cannot think of a scenario leading to the event. "The credibility of a scenario to a subject seems to depend more on the coherence with which its author has spun the tale than on its intrinsically 'logical' probability of occurrence."

In the same paper, Spetzler and Stael von Holstein describe methodologies for coping, at least in part, with bias tendencies and basic modes of judgment. A paper by Stael von Holstein (1974) describes procedures for encoding probability distributions from groups. Other useful references on multiple attribute probability assessments are Huber (1974), Fischer (1974), Schlaifer (1967) von Winterfeldt and Fischer (1973), and Kahneman and Tversky (1973).

Where objective probabilities can be compared with subjective estimates of probability, the most common finding in the literature is that naive subjects tend to overestimate the frequency of rare (improbable) events and under-estimate the frequency of extremely frequent events. The combination of these error tendencies results in an unduly restricted range of variation in probability estimates. Edwards, who has extensively investigated this pheno-menon, and other researchers point out that many persons can be rather quickly trained to get outside this restricted range. The training consists of exposure to the equivalent of empirical data tables that relate a pattern of outcomes to empirically (or logically) derived probabilities. Schum (1970) quotes Edwards at proposing that, because the human is not a very good numerical calculator, a computer should calculate probabilities if frequency distributions of various classes of event are actually known. But because the introduction of an innovation generates a relatively unique pattern of events, historical records uninterpreted by humans would have little direct applicability for determining prior probabilities. Edwards' recommendation is therefore limited to computational and aggregational operations on subjectively developed probability estimates.

In real life, a major difficulty in improving one's ability to estimate proba-
bilities in a given context is the long time between making the estimate and
getting feedback about the outcome that really happened. Goldberg (1968) cites
the results of studies on feedback of this kind in diagnostic situations: he
interpreted the small amount of improvement in estimating to the extended delays.
We have the same problem in assessing innovations.

Huber (1974) proposes three ways of increasing the accuracy of subjective judg-
ments of probabilities. One is to train the subjects, such as by tutoring in
statistics, discussion about potential biases and, where possible, giving feed-
back about accuracy of the judgments. A second approach is to aggregate esti-
mates from several estimators, in conditions where they are independent of each
other and also in conditions where they discuss the issues. In aggregating
estimates, little seems to be gained from weighting procedures, according to
Huber. The third approach is procedural. It consists primarily in consistency
checking of an estimate either by conversation with the "consultant" as done
by Huber, or by alternative modes of eliciting the estimate such as proposed
by Spetzler and Stael von Holstein (1972).

In the ETAM context, a fourth recommendation may be offered. That is to give
the assessing group, who will be making estimates of probabilities of various
outcomes, comprehensive information about the variables and conditions that
influence outcomes in given directions. But note that this recommendation is
explicit in Task 3 in the ETAM procedures where experts estimate the size of
the major obstacles and/or facilitations to successful implementation of an in-
novation. This might be called the analytic phase of information gathering and
evaluating that should precede the estimate of probability of given projected
outcomes. The more clearly the assessor can visualize the pattern of events
associated with predictions, the more valid and reliable we should expect the
probability estimates to be. And the clearer the assessor's implicit picture
of chances for success or failure, the less likely will one procedure differ
from others in eliciting the overt estimate.

There are three commonly used procedures for obtaining probability estimates.
One is that of direct estimation such as: outcome A has a 70% chance of
occurring. A second method is that of estimating odds, such as by saying that
one would take a bet of 7 to 3 that A will happen. This kind of estimating
of probabilities can be difficult unless the estimator has extensive experience
with the technique. A third method is an estimate of chances on proportional
probabilities such as in saying that one would take chances of 7 out of 10
that event A will occur.

DIFFERENTIAL PROBABILITIES OF BENEFITS IN DIFFERENT OUTCOMES

The procedures described in Section II assume that each projected outcome has
a benefit package: if the outcome happens, the entire benefit package happens.
The probability is attached to the entire projected outcome. Note that the
probabilities for a given outcome are determined from the assessment of
environmental influences with which the innovation is expected to cope. The
estimates are developed in ETAM Task 3, 4 and 6.

The assumption of equal probability for all elements in a benefit package may not be applicable. That is, the assessor may believe that individual benefits within one outcome could have differential probabilities of occurrence. The ETAM decision tree procedures permit expansion from a single predicted package of benefits to a variety of branches, each with a different package of benefit probabilities. The branches from a given outcome node may consist of multiple packages of varying estimated probability, where each package contains a different assortment of subjective probabilities of being achieved for each benefit. The branches of a decision tree can proliferate to any degree of outcome specificity with respect to component probabilities. The proliferation can add more information to the decision. The warning, however, is that if the assessor goes beyond a procedural level where he is able to contribute information to the problem, he will almost certainly be contributing noise (random variance and unreliability or irrelevance) to the picture.

SENSITIVITY ANALYSIS AND PREDICTED DISPERSIONS ON AN ATTRIBUTE

The ETAM procedures for assessing innovations, Task 5, call for the assessor to make a best estimate of the magnitude of incremental change to be expected of each benefit attribute if the innovation is adopted and implemented into the training/operational systems. The assessor is also asked to estimate the incremental improvement in the benefit attribute if it is implemented and accepted "under the most favorable conditions" and the improvement level to be achieved "under the least favorable conditions."

This range across the mean or median estimate (i.e., the most likely level of the outcome) is the range from an optimistic estimate to a pessimistic estimate of a benefit level. This range is assumed to be a segment of a beta distribution where: (see Battersby, pp. 170-175)

Standard deviation = (optimistic estimate minus pessimistic estimate)/6

In effect this implies that the optimistic to pessimistic limits define most of the distribution of benefit levels that are reasonable to occur as an outcome for a given path of action. It is not the purpose here to challenge or justify the logical assumptions underlying the calculation of this distribution. Rather, the concern here is with the psychological meaning and justification of the optimistic and pessimistic boundary limits. The magnitude of this range will signify two confounded factors in the assessor's thinking. One is his perception and estimate of the random variability in the operational world in which he is projecting the effects of the innovation. This projection will be based both on the data known to him about the innovation's effects in pilot tryouts and perhaps from less direct sources. He will have examined also the influences arrayed for and against the successful implementation of the innovation.

The second influence on the range between his optimistic and pessimistic estimates will be his own degree of uncertainty (or sense of ignorance) about the outcome even if the innovation is "successfully" implemented in the General sense. From the standpoint of making a decision, however, it makes little difference whether the range of variation comes from one or the other source. In either or both cases, this range is an appropriate reference from which to calculate the sensitivity of the total value of a projected outcome to a given benefit. If the benefit is a dominant one among benefits

(either positive or negative) a large range of uncertainty about its incremental
value logically introduces a higher level of risk into the decision. The
increased risk may be sufficient to decrease the attractiveness of an acceptance
decision to a point of rejection, or to a point that may justify the collection
of further data on this benefit variable through empirical or other investigation.

The logical transformations of these estimates, as referenced to the "most
expected" or median increment of change on the benefit variable, might better use
an assumption of skewness in the distribution around the "best estimate." It
seems reasonable that in many cases, a pessimistic estimate would be further
from the "best estimate" than the optimistic estimate would be. Regression
to a state of "no change" seems generally to have more factors working for it,
than does an increase above some best expected change level. The realistic
assessor may be aware of these dynamics, and reflect this awareness in his
judgments.

Whether an assumption of skewness rather than of symmetry of pessimism and
optimism around the best estimate point would make any practical difference
in making decisions is moot. If this consideration is like many other attempts
at refinements in scaling and decision models, it will make little or no
practical difference.

It does seem reasonable to give the assessor a picture of how his judgments of
the range set by his optimistic and pessimistic limits will be interpreted and
used by the internal program of the system. In effect, he should learn that
his "pessimistic point" means that he expected there are only 1½ chances in a
thousand that the benefit value would go below this level if the action path
to this outcome were followed. Here also, some exposure to elementary statistics
would give the assessor a sense of the scale that will be used for interpreting
his actions. Another alternative is to choose different descriptions for
these points. Example: "What is the worst level of the benefit attribute
that could reasonably happen if the innovation were successfully implemented
and accepted?"

We should not expect reliable judgments--or very useful ones--if we present
the assessor with the problem of locating the lowest point on the benefit
continuum that he expects has about 1 chance in a thousand of being achieved.
This is the kind of judgment that humans do poorly, and inevitably make
regressively. But interactive computer procedures can let the assessor make
a semi-qualitative judgment and then tell him how that judgment will be
interpreted and applied quantitatively. He may then revise his qualitative
judgment in the light of the "meaning" that has been shown to him.

CONCLUSION

If the lay reader enters into utility and decision theory by first reading
theoretical papers by writers such as Fischer or Fishburn, he may be over-
whelmed by the complexities of the subject matter and the array of alternative
"models" from which to choose. If, however, the reader first enters the field
by reading the reports by practitioners working with clients in the field
such as Schlaifer and Huber, he becomes impressed with the simplicity and
directness of scaling procedures for probabilities and utilities. If he also

reads the critiques of empirical studies comparing techniques for scaling, he finds that the simplest methods, based on the simplest assumptions, work just about as well--or equally as well--as the more sophisticated methods and calculational models. Even when fundamental assumptions such as linearity and independence of attributes are violated, there does not usually seem to be great differences in practical results.

The interactive computer would facilitate the use of alternative methods on the same problem in order to determine consistency of judgments, and-- where differences appear--to help in reconciling them. The computer also can make it easy to introduce changes, firm or tentative, and recalculate the results both proximal and remote in the entire decison problem. Finally, the capability of an adequately programmed computer that enables the user to selectively retrieve operational facts and references that support his generalized judgments, as well as the series of judgments he has already made, should greatly increase the actual information that is <u>used</u> by the assessor and decision maker in assessing and deciding. In this regard, both the computer and scaling procedures, including decomposition techniques, are assets for training and for exercising and supporting the informed human with a problem to solve. Finally, the computer can interject warnings about overzealous acceptance of the outputs generated through it or by it.

This conclusion is especially applicable to assessments and decisions about innovations in large systems such as personnel, training and operations where the potential effects of change can have wide ranging patterns of influence and counter-influence. The structuring of assessment and decision processes is a desirable product in itself. So is the capacity to communicate explicit actions about explicit information either objective or subjective. If a methodology can help the human to formulate and objectify what often are vague value systems about the world in which he makes decisions, so much the better, and especially if the methods are unobtrusive and reasonable to the user. We have applied these criteria in formulating the procedural recommendations in Section II.

This section is best concluded with quotations from two eminent researchers with many years of study into behavioral and mathematical models in decision making. Edwards (1975) terminates a report that summarizes a broad and in- tensive study of behavioral decision making as follows:

> "...In decision analysis, structuring the problem and processing the information are of primary importance, while eliciting proba- bilities and utilities are of secondary importance...Research on the merits of information sources, on optimization of information processing, and on formulation of decision problems is more impor- tant than work on precise elicitation and optimization procedures."

This statement should also apply to applications, and supports the procedural care in ETAM I for the generating and analysis of benefit and cost attributes and values, as well as expected impediments and liabilities, that may apply to any decision about any innovation. In a similar vein, von Winterfeldt (1975) is generally pessimistic about the refinements and complexities introduced by mathematicians into utility theory, and acknowledges the importance of simplicity in the practical application, even though various assumptions may be uncertain.

In concluding his survey report (p. 69) he advises:

> "But if one wants to find simple, quick, realistic methods that
> produce little error, one will have to look outside the realm
> of theoretically feasible methods. Probably the most reasonable
> methods of this sort are magnitude estimation methods such as
> direct rating, direct judgment of utility differences, direct
> ratio assessment of weights, direct assessment of probabilities
> and utilities, etc. Clearly they are uncomplicated. All the
> assessor has to do is to quantify his judgment on a numerical
> scale."

On the next page he remarks:

> "But it appears that there is at least a reasonable tradeoff
> between quick, simple and realistic direct estimation methods
> that are not formally justified in the model context and the
> somewhat clumsy feasible indifference methods."

We have heeded this advice, which emerges from several decades of dedi-
cated inquiry.

APPENDIX I

It is often meaningful and convenient to use an abstract qualitative concept to denote a variety of implications that sum into something good (such as increases in "motivation") or into something bad (like "delinquency" or dependency on others"). The qualitative expression may be used by a specialist who is educated in its denotational operations; but sometimes this is not the case. In any event, there are a number of terms frequently used in the training and operational environments that deserve some denotational handles for grasping by assessors and decision makers.

The following are examples whereby a complex qualitative term is decomposed into operational, observable and measureable or countable, terms. The decomposition cannot be exhaustive, but can only offer examples of operational reference. Decomposition of abstract variables into denotable operations often reveals heavy overlap of meanings.

This list should be integrated with the Range of Effect terms contained in the topical outline to ETAM, Task 5. The following lists also indicate that benefit/liability variables other than directly costable ones can be associated with a benefit variable.

SUPPLEMENTAL BENEFITS/LIABILITIES OF <u>ATTRITION RATE</u>

These are some variables associated with attrition rate that are not directly costable:

    Lower acceptance levels for a course of training
    Greater availability of trained personnel
    Greater predictability of training outcome in terms of number of graduates
    Greater flexibility in other training variables (holding acceptable output
        criteria constant)--such as preparation of training; administering
        of training
    Greater confidence in accepting the training by the potential candidates
    Greater range among individual competences on the job (if the reduced
        attrition rate is because of special treatment of only the poorer
        students)
    Liability: Potential loss of flexibility in assignements--because of the
        lower aptitude level or because of more specifically directed skills.

SUPPLEMENTAL BENEFITS/LIABILITIES: SHORTER TRAINING TIME

    Increases skill level--if old training time is maintained
    Reduced attrition--if innovation were applied using old training time
    Lower cost of instructional facilities
    Extended time of incumbents for operations during limited service
        durations
    Shorter preparation time for manning new systems (faster preparedness)
    Greater flexibility in weapon system development/procurement
    Greater temporal flexibility in manpower deployment and commitment
    Faster responsiveness to specific skill demand in operations because of:

    a. Less overcommitments in training or operations

    b. Less waste from uncertainty of demand/resource balance

More flexibility in deployment of instructional staff
High level instructional skills and motivations required of instruc-
tional staff: this is a liability

SUPPLEMENTAL BENEFITS/LIABILITIES: APTITUDE REQUIREMENTS

Note that Aptitude Requirements are non-costable in any strict sense until an
empirical model of aptitude availability and demand exists, with dollar values
attached to the demand.

Tradeoff principle: As an aptitude level is reduced by some increment, a
    disproportionately larger increment of potentially acceptable indivi-
    duals is created. This is especially the case where high levels of any
    kind of aptitude is the starting point. Where more than one non-
    correlated aptitude is reduced at the same time, the consequent increase
    in eligible persons in a general manpower pool can become enormous.

Other factors constant, reduction in an aptitude level requirement may be
traded off against:

1. Reduction in attrition rate

        or

2. Reduction in training time

    Note: Reduction in training time will be affected less than attri-
    tion level because the distribution of aptitudes above the old cutoff
    aptitude score may not be affected by the innovation.

3. Reduction of critical aptitudes increases the likelihood that any
    available manpower pool will contain enough qualified individuals
    to fill given class quotas.

4. Reduction of aptitude levels increases the likelihood that diverse
    aptitude requirements can be met from a given heterogeneity in a
    manpower pool.

Liability: A minimum aptitude for achieving a given level of formal training
    may be inadequate for on-the-job requirements, or for changes in on-the-
    job requirements.

    Note, however, that by the same argument, the aptitude level required to
    pass training requirements may be higher in some attributes than required
    by on-the-job demands.

    Reductions in aptitudes associated with general intelligence can reduce
    the extent to whcih the selected individuals are capable of self-instruc-
    tion and "learning from experience" in cognitive tasks such as decision
    making, planning, organizing, and constructing.

SUPPLEMENTAL BENEFITS/LIABILITIES: INCREASED <u>JOB-RELATED</u> <u>SKILL</u> <u>LEVEL</u>, OR <u>SKILL</u> <u>RANGE</u>

Most of the following factors are qualifying conditions for assessing the value of increased job-related skill levels and ranges.

1. The job may provide little opportunity for applying increased level or range of skills, or place little value on them when applied.

    Skills rarely used on the job tend quickly to degrade, although they may be relearned rather quickly given proper opportunity for controlled practice.

2. Reference knowledges intended to support job skills degrade especially quickly if not periodically refreshed by controlled practice conditions. If the knowledge is relatively abstract, many operators may even forget that they have ever been exposed to it.

3. <u>Except</u> for achieving certain critical levels of skill learning, even fairly large incremental increases in skill level acquired in formal training will often be largely diluted in on-the-job performance by motivational, environmental and operational factors anticipated in formal training.

    Note that this observation tends to apply more to procedural tasks than to teaching tasks that require anticipatory behavior about situations and about the dynamic properties of equipment used (e.g., piloting aircraft, operating vehicles, manual aiming of guns.)

    Secondary or auxiliary skill may enable the operator to take over a task B when the primary operator for B is unavailable for performing it. Some assessment should be made of the probability of the take-over being required and practical in the operational scene, the availability of operator A being free to perform task B, and the importance of B to the type of mission in which task B occurs. This combination of factors, probability of opportunity and task importance, should be balanced against the cost of the training and the probability that the incumbent will retain the skill/knowledge when the opportunity arises for using it.

4. See also the ETAM list of job benefits potentially arising from increased skill level training:

        Task 05, Subtask 15 SELECT RELEVANT STAGES OF LEARNING
        Task 05, Subtask 23 ASSESS ON-THE-JOB BENEFITS

SOME MEASURABLE FACTORS: MOTIVATION TO LEARN

Examples of operations associated with increases of motivation to learn, up to a limit where further increases serve as counterproductive stress, are:

1. Better retention measures

2. Capacity for more protracted periods of study

3. Recognizes specific deficiencies and directs self to overcoming them

4. Resists fatigue longer; recovers more quickly

5. Keeps on trying after greater frequency of frustrations

6. Keeps on practicing for higher criterion levels

7. Uses more resources available for learning

8. Directs himself to study or practice independent of social pressures

9. Overcomes environmental impediments in order to practice in learning a knowledge or skill

10. Overcomes psychological distractions such as from competing interests, incentives, attitudes and preferences, showing this by study and practice.

11. Finds opportunities to apply directly or indirectly what is learned

12. Performance shows more rapid warm-up effect in practice sessions.

SOME MEASUREABLE FACTORS: INITIATIVE IN LEARNING/PERFORMANCE

Note: This factor has overlap with operations that are identified under "MOTIVATION." Essentially, initiative is the directing of oneself rather than being directed by others, and by individual action as contrasted with rule-directed action. Whether or not the outcome of an initiative is good or bad is not germane to the fact that an initiative has been taken. Initiative is shown in the following examples:

1. Noting and reporting events not covered by rules

2. Taking task-related or system-related action not covered by rules

3. Disobeying or disregarding a procedure or directive, presumably on the basis of common sense, or special knowledge, or because of special conditions

4. Performing a task as the occasion warrants without being told to do so

5. Performing auxiliary or supportive activities beyond the minimum prescribed requirements

6. Performing tasks or task action not prescribed for the individual to do

7. Applying knowledge to action performed in a novel situation.

SOME MEASURABLE FACTORS: PROFESSIONALISM

This usually connotes a combination of motivations and knowledge and skills
that exceed some minimum standard of requirements for a job or task, and in-
cludes a continuously higher level of personal service and aspiration. It is
manifest by, for example:

1. Finding opportunities for further formal and informal learning--
   this applies especially to updating his knowledge and skills

2. Responding to requests for service when not formally required to
   do so

3. Carrying through a task until evidence confirms that the work has
   been successful

4. Acknowledging unsuccessful task performance if accountability can
   lead to a remedy or to identification of a responsibility

5. Using auxiliary resources in getting a service performed, including
   the skills of other people

6. Increasing the range of his knowledge of technologies that opera-
   tionally interfaces with his services

7. Giving effort appropriate to the demand for services rather than
   as a standard set of operations

8. Applying ethics--in terms of right or wrong--in situations where a
   personal cost may be involved

9. Acknowledging ignorance or incompetence where a demand for service
   exceeds capability and some other available resource could better
   perform the service

10. Teaching others his knowledge and skills formally or informally.

BIBLIOGRAPHY

Battersby, A., 1964. Network Analysis for Planning and Scheduling. Macmillan and Company Limited.

Baumol, W. J., 1965. Economic Theory and Operations Analysis. 2nd Edition. Englewood Cliffs, N.Y. Prentice-Hall.

Chapman, Lid and J. P. Chapman, 1967. Genesis of Popular But Erroneous Psychodiagnostic Observations. Journal of Abnormal Psychology. Vol. 72, pp. 193-204.

Eckenrode, R. T., 1965. Weighting Multiple Criteria. Management Sciences. Vol. 12, No. 3.

Edwards. W., 1961. Behavioral Decision Theory. Annual Review of Psychology. 12, 473-498.

Edwards, W., 1963. Probabilistic Information Processing by Man, Machines and Man-Machine Systems. Santa Monica, Calif. System Development Corp. TM-1418/000/01., August.

Edwards, W., 1975. Research on the Technology Inference and Decision. Final Technical Report. SSRI Report 75-10.

Fisher, G. W., 1972. Four Methods for Assessing Multi-Attribute Utilities: An Experimental Validation. ONR Engineering Psychology Programs. NR197-014.

Fishburn, P. C., 1966. Decision Under Uncertainty: An Introductory Exposition. Journal of Industrial Engineering 27(7).

Fishburn, P. C., 1967. Methods of Estimating Additive Utilities. Management Sciences. Vol. 13 No. 7.

Freedy, A., et al., 1974. Adaptive Computer Aiding in Dynamic Decision Processes Part I. Adaptive Decision Models and Dynamic Utility Estimation. Technical Report PTR-1016-74-5 (I), Perceptronics, Inc., Prepared for the Office of Naval Research under Contract N00014-73-C-0286.

Giaque, W. C., 1975. Decision Making in the Naval Education and Training System, TAEG Report No. 27, Training Analysis and Evaluation Group, Orlando, FL.

Goldberg, L. R., 1968. Simple Models or Simple Processes? American Psychologist Vol. 23, pp 483-496.

Guilford, J. P., 1954. Psychometric Methods, N. Y. McGraw-Hill.

Gustafson, D. H., G. K. Pai and G. C. Kramer, 1971. A "Weighted Aggregate" Approach to R & D Selection, AIIE Transactions, Vol. III, No. 1.

Halpin, S. M. and J. A. Thornberry and S. Streufert, 1973. The Credibility of Computer Estimates in a Simple Decision Making Task. ONR Technical Report #5. ONR Organizational Effectiveness Research Programs, Code 2452.

Nav train Math mods

Comp prog. corry semul

Handbook for Decision Analysis, 1973. A user-oriented document prepared by
    Decision and Design, Inc., for the Advanced Research Projects Agency and
    Office of Naval Research under Contract N00014-73C-0149.

Huber, G. P., 1974. Methods for Quantifying Subjective Probabilities and
    Multi-Attribute Utilities. The Journal for the American Institute for
    Decision Sciences. Vol. 5, No. 3.

Huber, G. P., 1974. Multi-Attribute Utility Models: A Review of Field
    and Field-Like Studies. Management Sciences. Vol. 20, No. 10. June.

Kahneman, D. and A. Tversky, 1973. On the Psychology of Prediction.
    Psychological Review. Vol. 80, No. 4. (July).

Miller, G. A., 1956. The Magical Number Seven Plus or Minus Two: Some
    Limits on Our Capacity for Processing Information. Psychological Review
    Vol. 63, 81-97.

Miller, Robert B. and Larry R. Duffy, 1975. Design of Training Systems, Phase
    IIA Final Report, TAEG Report No. 12-3. Training Analysis and Evaluation
    Group, Orlando, FL.   A015295

Murphy, A. H. and R. L. Winkler, 1973. Subjective Probability Forecasting of
    Temperature: Some Experimental Results. Proceedings of the Third
    Conference on Statistics in Atmospheric Science. Amer. Meteorological
    Society. Boulder, Colorado, June.

Peterson, C. R., 1971. Judgments of Probability and Utility for Decision
    Making. First Annual Report. DMR Contract No. N00014-67-A-0181-0034.
    NR 197-014.

Peterson, C. and L. Beach, 1967. Man as an Intuitive Statistician.
    Psychological Bulletin, July, 68 (1), 29-46.

Schlaiffer, R., 1967. Analysis of Decisions Under Uncertainty, Vol. 1.
    N. Y. McGraw-Hill.

Schum, D. A., 1970. Behavioral Decision Theory and Man-Machine Systems.
    In DeGreene (Ed.) Systems Psychology, N. Y. McGraw-Hill.

Slovic, P., 1966. Value as a Determiner of Subjective Probability. IEEE
    Transactions on Human Factors in Electronics, March, HFE-7, 22-28.

Slovic, P. and S. Lichtenstein, Nov. 1971, Comparison of Bayesian and
    Regression Approaches to the Study of Information Processing Judgment.
    Organizational Behavior and Human Performance, Vol. 6, No. 6, pp. 649-744.

Spetzler, C. S. and C. S. Stael von Holstein, 1972. Probability Encoding in
    Decision Analysis. Paper presented at the ORSA-TIMS-AIEE Joint National
    Meeting, Atlantic City, N.J. (November).

Stael von Holstein, C. S., 1974. Procedures for Encoding Probability
    Distributions from Groups. Menlo Park, Calif. Stanford Research
    Institute Project Memorandum, SRI Project 2753-2, May.

v. Winterfeldt, D., 1971. Multi-Attribute Utility Theory: Theoretical Background and an Experimental Validation. Diploma Thesis, University of Hamburg (Germany).

v. Winterfeldt, D. and G. W. Fischer, 1973. Multi-Attribute Utility Theory: Models and Assessment Procedures. Invited review paper presented at the Fourth Conference on Subjective Probability, Utility and Decision Making, Rome, Sept. 3-6.

v. Winterfeldt, D., 1975. An Overview, Integration and Evaluation of Utility Theory for Decision Analysis, Technical Report, sponsored by DARPA, SSRI Research Report 75-9.

Naval Training Equipment Center, Orlando, Fla.

TAEG REPORT NO. 32   UNCLASSIFIED
                      AD _____

DESIGN OF TRAINING SYSTEMS
THE DEVELOPMENT OF SCALING PROCEDURES
1 Volume, 86 p. 11 illus.

This single volume final report expands the knowledge and procedural base for making decisions where the potential outcomes of alternative decision actions are characterized by multiple dimensions of value. This extends and heavily relies upon the work previously reported in the Phase II-A Report.

The report includes extensive references to the current literature on multi-attribute utility estimating in discussing both the theoretic as well as the practical considerations faced by assessors and decision makers in structuring data for decision making purposes. Finally, a detailed set of procedures is presented in Section II; it is intended to provide design guidance in developing an interactive computer-based program of Educational Technology Assessment Model (ETAM).

Training Mgmt. Systems
Computer Based Models
Educational Technology
Training Technology
Decision Theory
Utility Theory
Subjective Estimating

IBM Corporation
Federal Systems Division
Civil Space Systems

Contract No.
N61339-73-C-0097