Central Intelligence Agency

Washington, D.C. 20505

3 July 2001

Mr. John Greenewald, Jr.

Reference: F-2001-01030

Dear Mr. Greenewald:

This is to acknowledge receipt of your check number 1104 dated 20 June 2001 in the amount of $24.90. In response to your Freedom of Information Act request dated 24 April 2001, we have provided the enclosed document, *Handbook of Problem Solving Techniques for Intelligence Analysts*.

Thank you for your payment. We have credited your account in full.

Sincerely,

Kathryn I. Dyer
Information and Privacy Coordinator

Enclosures

HANDBOOK of

PROBLEM-SOLVING TECHNIQUES

for Intelligence Analysts

Working Draft:

Current as of November 21, 1986

PROBLEM SOLVING TECHNIQUES

PRECIS

Analysts are professional problem solvers. The extent to which they effectively use good problem-solving techniques directly impacts the quality and usefulness of resulting intelligence products. Nonetheless, until recently few analysts have received formal training in problem-solving techniques that can help them more effectively and efficiently manipulate data to extract policy relevant information for presentation to consumers in finished intelligence products and briefings.

This Handbook provides (in the Introduction) an overview of the problem-solving process as it applies to intelligence analysis. Subsequent sections of the Handbook describe selected techniques potentially useful to different phases of the overall problem-solving process.

In general, problem solving consists of a sequential series of activities. This Handbook categorizes them into three major phases.

The Planning Phase. This Phase concerns getting organized to solve a problem. It includes methodological approaches to such things as problem structuring, selecting appropriate analytic techniques, and selecting or redefining problem statements.

The Analysis Phase. This phase involves using the techniques identified as appropriate (whether they be heavily quantitative or predominantly intuitive techniques) to solve the problem.

The Production Phase. This phase includes the activities involved in preparing final/formal briefings and written products. This includes such things as briefing and writing techniques, formal logic, and suggestions for making useful graphics.

This Handbook is not designed to be read cover to cover like a textbook. The beginning of the section on the Problem-Solving Process is recommended as general reading, and the section on selecting appropriate analytic techniques is generally useful also. The remainder of the Handbook is intended to be a readily useable reference manual. It is organized to be used like a cookbook, so that if help is needed with some specific portion of the problem-solving process, a sampling of potentially useful techniques can be looked up and tried.

PROBLEM SOLVING TECHNIQUES

TABLE OF CONTENTS

PROBLEM SOLVING TECHNIQUES

PLANNING ACTIVITIES

PROBLEM SOLVING TECHNIQUES

ANALYSIS ACTIVITIES

PROBLEM SOLVING TECHNIQUES

PROBLEM SOLVING TECHNIQUES

PROBLEM SOLVING TECHNIQUES

PROBLEM SOLVING TECHNIQUES

PRODUCTION ACTIVITIES

PROBLEM SOLVING TECHNIQUES

LIST OF FIGURES

page

PROBLEM SOLVING TECHNIQUES

PROBLEM SOLVING TECHNIQUES

LIST OF TABLES

page

PROBLEM SOLVING TECHNIQUES

"...seek simplicity and distrust it."[1]

---

[1] Anatol Rapoport, "The Search for Simplicity," The Relevance of General Systems Theory, Ervin Laszlo, ed., New York: George Braziller, 1972, p.

INTRODUCTION
Information and Intelligence Analysis

FOREWORD

INFORMATION AND INTELLIGENCE ANALYSIS

The real world, about which we make intelligence assessments, is composed of matter and energy. The information that must be provided to policy makers is carried via the same medium, but must have a radically different content than physics equations for it to be useful. This implies that the raw data (information) obtained from the real world must be transformed dramatically from the form in which we receive it initially; for collection systems (both human and technical) are basically energy transducers that provide samples of physical manifestations.

As analysts, we are professional problem solvers. We do not incur the managerial problem-solving role, where problem solutions must be implemented as well as identified. Rather, as the job title implies, analysts specialize in problem analysis. Thus, the job of an analyst largely involves manipulating data to extract policy relevant information for presentation to consumers in finished intelligence products.

The data chosen for manipulation is a function of existing intelligence problems and our ability as analysts to identify and use available information effectively. The finished intelligence products typically belong to one of the four taxonomic categories that Table 1 describes. Each of the four categories of finished intelligence products has its unique set of goals, customers, and analytical tasks. Hence, differing analytical methods tend to be associated with each. But, the end-product of analytical labors is, in large part, a function of the methods used. That is, the methods chosen to collect, organize, manipulate, interpret, and present information to a large extent determine (more precisely predetermine) the kind of conclusions that will be reached. Thus, we should, as analysts, try to optimize our selection of analytical strategies.

Problem-solving research has led to the conclusion that several generic types of problems can be identified, and that specific categories of techniques that best match the problem-solving needs and constraints of each problem type, can be specified. Hence, developing a familiarity with multiple approaches to analysis, and learning guidelines for identifying generic problem types and for selecting an approach(s) suitable to the problem at hand, would seem both advisable and useful.

This handbook develops a problem-solving approach generally applicable to all categories of analysis, irrespective of the substantive area being addressed. The section on the Problem-Solving Process begins on page 7. A later section, that starts on page 50, presents a strategy for categorizing

---

[1] This taxonomy was developed by William Schultheis (Chief Scientist, ORD/PATG) for the Intelligence Production Laboratory.

PROBLEM SOLVING TECHNIQUES
Information and Intelligence Analysis

TABLE 1 -- TAXONOMY OF INTELLIGENCE PRODUCTS

CATEGORY[a] Characteristics	Customers/Uses
CURRENT INTELLIGENCE Mainly descriptive. The analyst tries to create for the policy maker a coherent picture of a usually dynamic situation. Provides information updates to supplement long-term studies. Projections, if present, are short term. Is data driven with relatively little user feedback. Seldom uses formal methods. Often criticized for lack of depth and failure to predict the unexpected.	1. Policy Makers. To support short-term decision making. 2. Technical and policy analysts. To keep tabs on what's going on.
IN-DEPTH ANALYSIS Both descriptive and evaluative. Focused on a subject or situation -- usually from the standpoint of a single discipline. Products are scheduled or triggered by availability of information. Use of formal analytical methods extensive in some technical and economic areas to produce implicit or explicit projections. Problems include lack of focus on customer needs, lack of an interdisciplinary perspective, and lack of timeliness (contents out-of-date before report gets to user).	1. Technical and policy analysts. 2. Decision support staffs. (Used to obtain quite specific data -- often from rather general products)
ESTIMATIVE Focused on specific issues and broad questions. Perspective intended to be interdisciplinary. "Big Picture" approach with tables, figures, statistics where appropriate. Projections are made within a formalized framework particular to the subject matter. Interagency interaction involved (can become heated and politically motivated). Institutionalized dissent (footnotes). Participants worry when they agree. A major problem is now to sort out and evaluate alternative explanations and scenarios.	1. Strategic Planners Used for major scheduled decisions (e.g., budget) and policy evaluation.
QUICK REACTION (Not included in conventional taxonomies, but may be as important to users as any or all of the other categories.) Intelligence producer is told the specific user's problem and responds directly. Products can be memos, staff papers, briefings, phone calls, meetings, or (rarely) formal products. Reply is based on current knowledge (no time to collect -- rarely time to accomplish more than a hasty, subjective	1. Anyone with an important short-term need.

analysis). Users tend to be satisfied. Big problem is disruption to other activities (sometimes there is a special staff for such activities). Problems include determining where in the organization to look for the best information, and bureaucratic involvement in the response (what's the current format for the director's signature?).

problems into categories that can be used to select particularly appropriate techniques. The remainder of the handbook is primarily a cookbook, providing an introduction to a variety of potentially useful techniques. This broad-spectrum coverage of multiple techniques and technique types (from highly mathematical to extremely intuitive) provides a feeling for the vast variety of techniques that are available for analytical use, and concurrently gives a perspective on the effect that each technique will have on conclusions reached via its use.

A reader certainly is not expected to begin using every technique presented here. But hopefully a few that seem effective will be found; and concurrently, an understanding or appreciation of the techniques used by other analysts (your colleagues) will be gained. Equally important, it is hoped that the reader will gain some appreciation for the strategies that can be used when selecting a problem-solving approach.

PERSPECTIVE ON PROBLEM SOLVING TECHNIQUES

As suggested above, this handbook concentrates on those aspects of analytical skills that can be approached methodologically. Its intent is to provide exposure to a broad spectrum of techniques. As you review this handbook try the techniques out on a few trial problems, (even if they are trivial ones). You then can compare, contrast, and chose among the techniques from a base of at least limited experience.

The techniques are organized according to the where they most frequently fit into the generic problem-solving sequence used in this handbook: namely, as a Pre-Evaluation tool, as a tool used during the Evaluation Phase, or as one used for Presenting Conclusions. Some of the techniques are quite formal (structured), whereas others may seem so informal as to hardly be "techniques" at all. They vary also along the dimensions of Veridicality (subjective versus factual) and Complexity (simple versus complex). Each of the dimensions has advantages and disadvantages associated with its extremes. Table 2 gives some thoughts on this. The UPPERCASE factors indicate advantages, the lowercase factors, disadvantages.

No one technique is designed to address all types of analytical problems; so please do not look for or expect a methodological panacea. Do

TABLE 2 -- COMPARISON OF DIFFERING TECHNIQUE TYPES

COMPLEXITY

Simple	Complex
FAST	slow
EASY	hard
CHEAP	expensive
	confusing
non-enlightening	INSIGHT HELPFUL
narrow focus	COMPREHENSIVE
limited capacity	LARGE CAPACITY

STRUCTURE

Unstructured	Structured
FLEXIBLE	rigid (inhibits creativity)
potentially inconsistent	CONSISTENT
undocumented	DOCUMENTED
updating sophisticated	EASILY UPDATED

VERIDICALITY

Subjective	Veridical
LOW DATA REQUIREMENTS	requires hard data
soft results	CORRECT RESULTS

not become captive of a favored technique, nor overload yourself with so many techniques (or with so much detail about any of the more mathematically intricate techniques) that you fail to gain fundamental skills with at least a few.

Pick and choose among the techniques presented. Note the weak points and limitations as well as advantages of each. Every technique has weak points, if only in that it predetermines the type or limits the range of conclusions that can be obtained through its use. Search for the logic, assumptions and purposes of the methods presented. Question freely. Seek to learn and to internalize the methods.

INTRODUCTION
Perspective on Problem Solving Techniques

CATEGORIES OF PROBLEMS

Several schemes exist for placing problems into generic categories. VanGundy[3] discusses three types of problems:

1. Well-Structured Problems. Ones that have all the information needed to analyze a problem situation and to prescribe corrective (or optimal) solutions. Approaches to this type of problem can be intuitive if the problem is simple, and if complex, can be addressed by known algorithms (recipes). When approached using a known method (algorithm) a correct or optimal solution is a virtual certainty.

2. Semi-Structured Problems. Ones that have sufficient data to partially define the nature of the problem, but which leave some uncertainty as to the exact current state-of-nature, the exact desired goal, or how to best proceed. These types of problems are usually best approached heuristically. That is, by using rules of thumb or guidelines as how to proceed. Heuristic approaches do not guarantee a best solution, they simply increase the likelihood that a successful solution will be found.

3. Ill-Structured Problems. Ones that provide little or no clue as to how best to approach solving the problem. These kinds of problems require the development of solutions custom-made for the particular problem at hand. This often requires much improvising; hence creative problem solving techniques often are useful or necessary.

Most everyday problems fall into the first or second categories. Most intelligence problems fall into the second or third categories (the first time that the particular problem type is encountered), with the majority falling into the second. One usually is best off by approaching a problem with a congruently matched problem-solving technique (using creative problem solving techniques on a simple and obvious problem will waste time and resources, and using an intuitive or recipe approach on an ill-defined problem will rarely produce a useful or anywhere-close-to-optimal solution). We should try to be versatile (to match problem-solving techniques and resources to problem type and requirements), rather than treating every problem in exactly the same way (whether that way is a formalized methodological one or a hip-pocket, intuitive approach).

This handbook predominantly concentrates on presenting algorithmic, with a smattering of heuristic, approaches to problem solving. It is, nonetheless, relevant to addressing ill-structured problems requiring

----

[3] VanGundy, A.B. Techniques of Structured Problem Solving. New York: Van Nostrand Reinhold Company, 1981.

creative problem-solving techniques. Creative problem solving typically draws heavily upon standardized algorithmic (recipe) and heuristic techniques, which are simply reconfigured or mated in new or unusual ways. For instance, decision analytic techniques were originally a creative solution to the ill-structured problem of making decisions when conditions of uncertainty prevailed and when the value of possible decision outcomes could be assessed only subjectively. The creative solution that decision analysis represents is simply the combining--which had not been done before--of the standard and well understood algorithmic techniques of probability and utility theories. Einstein's theory of relativity can be viewed in much the same way--combining Newtonian physics with a constant speed of light. Therefore, think of the techniques presented here as tools from which you can build other (new and creative) tools; rather than a set of tools to which you should try to force-fit every bloody problem you encounter. Use the techniques to further expand your capabilities, not to define their limits.

INTRODUCTION

THE PROBLEM-SOLVING PROCESS

A GENERAL PROBLEM SOLVING APPROACH

Problem solving is a multi-stage process. Usually it is addressed from the perspective of a decision maker who has the power (and the responsibility) to use the problem-solving process to generate solutions (in the form of action plans) that will be implemented under his/her direction. Once implemented, the actions taken directly impact the real world and the problem area itself (hopefully with the result of eliminating the problem altogether, or at least reducing it in scope or severity). The problem-solving activities of an analyst differ in that it is the policy maker, not the analyst, who must choose an appropriate course of action and implement it. The analyst's charter is concerned with problem solving for the purpose of providing information to policy makers to aid their decision tasks. Nonetheless, a strong parallel can be found between standard problem-solving methods, and the problem-solving process specific to analytical functions.

Many problem-solving models have been proposed. We will use an approach that parallels one of the better known models--Simon's three-stage process of: intelligence, design, and choice.[4] In the intelligence stage of Simon's process, the problem is identified and information is gathered to formulate an appropriate problem definition. Simon's design stage consists of developing problem solutions, while his choice stage, selects and implements the solution alternatives.

As VanGundy[5] points out, Simon's three problem-solving stages are characterized by convergent and divergent processes. The convergent processes involve a narrowing down of information to some manageable unit; divergent acts are just the opposite: information is reorganized or generated to create an increased scope or breath of coverage. Simon's intelligence stage involves dealing with a large amount of information about a problem, analyzing it, and then focusing attention on a limited number of definitions so that a working statement of the problem can be achieved. The design stage involves a divergent process of generating many different ideas. Finally, the choice stage consists of reducing the pool of ideas to one or two that will be likely to solve the problem. The solution is then implemented and evaluated for its effect upon the problem.

VanGundy expands upon this model by treating each of Simon's stages as a problem-solving stage in minature. He divides each of the major problem-solving phases into substages of intelligence, design, and choice. A

---

[4] Simon, H.A. The New Science of Management Decision, rev. ed. Englewood Cliffs, N.J.: Prentice-Hall, 1977.

[5] VanGundy, A.B. Techniques of Structured Problem Solving. New York: Van Nostrand Reinhold Company, 1981.

parallel model will be developed here. It consists of the three problem-solving phases indicated in Figure 1. Entering the problem at the top of Figure 1 and moving downward through the problem solving process, each phase can be seen to consist of divergent, followed by convergent, activities, as is suggested by the initially expanding, then contracting boundaries drawn around the indicated activities. The first two phases, the Planning phase and the Analysis phase, are similar to the Simon intelligence and design stages, respectively. The third phase, Presenting Conclusions, differs from Simon's choice stage. The activity in this phase is directed toward "packaging" the results and interpretations of the analysis phase in a form suitable for conveying needed information to consumers. (In a sense, this could be viewed as an implementing stage--where the choice involves deciding on the best "information packaging" vehicle(s). But, to emphasize the absence of solution implementations that are designed to impact directly on the problem gap that was the subject of study--a major characteristic of Simon's model--the less active "Presenting Conclusions" lable will be used.) The contents of this handbook are organized into these three divisions.

The three-phase model that we shall follow in this text assumes that a problem situation is initiated by environmental stimuli (e.g., world events or a policy-maker's tasking). The actual problem solving activity begins with the recognition that a problem situation (e.g., an intelligence gap) exists. The activities involved in the problem recognition (or problem identification) process often are considered as a separate problem-solving stage, but, for simplicity, we shall incorporate it in our first phase.

Thus, the Planning phase is started with an intelligence substage wherein the intelligence gap or need is initially identified, and collection activities are initiated to acquire problem-relevant information. The design substage then is initiated in which alternative problem definitions (and redefinitions) are generated in combination with problem analysis to categorize the type of problem that is being addressed so that potentially relevant analysis techniques (methodological strategies) can be identified. A constraint analysis is performed also, to assess environmental factors (e.g., time, resource availability) that will impact subsequent problem-solving activities. The Planning phase ends with the choice substage wherein a working definition of the problem statement is selected, the problem-type to which the working definition belongs is identified, and the most appropriate analysis technique(s) is selected.

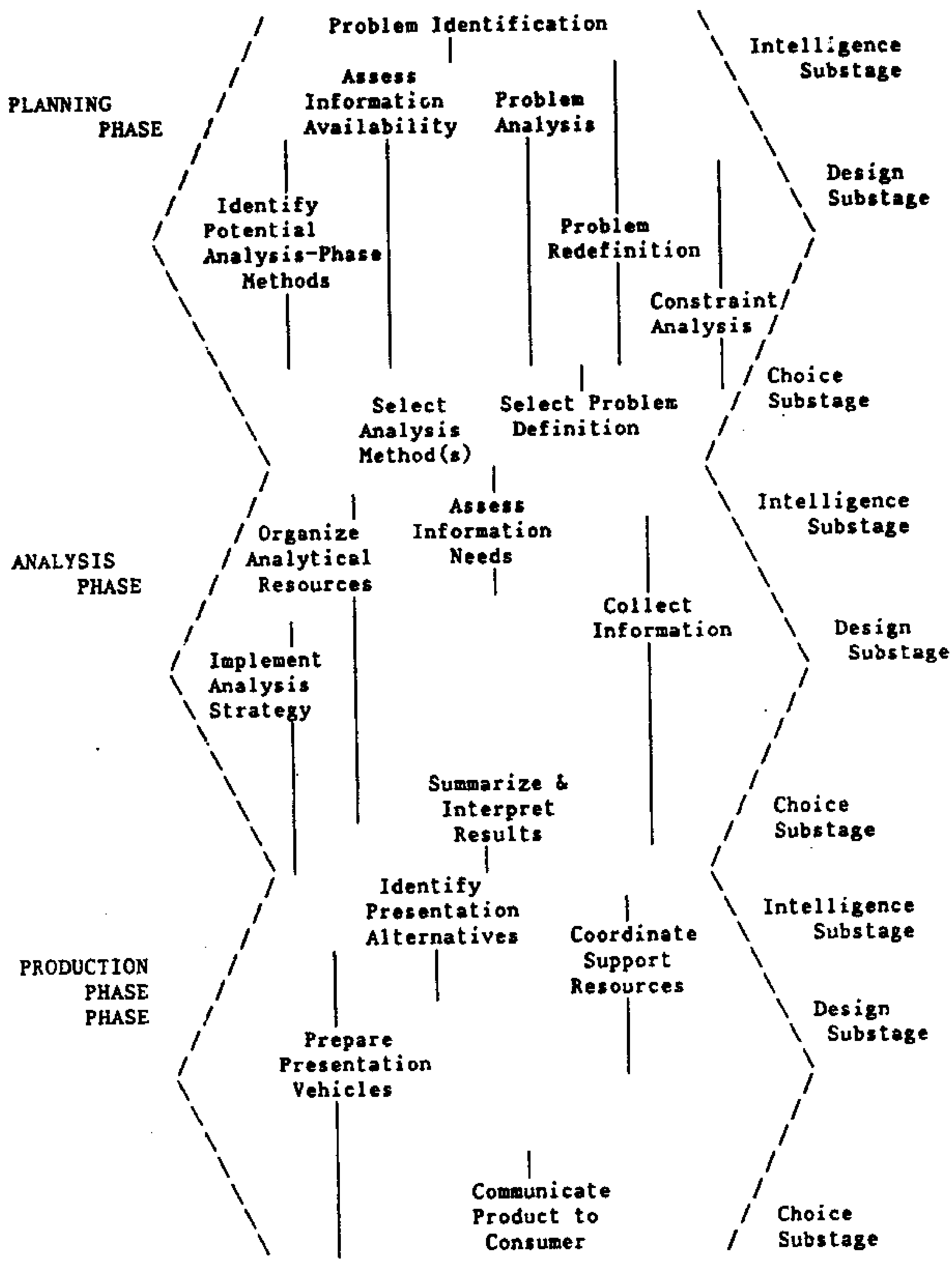
The intelligence substage of the Analysis phase begins with an assessment of the information needs and the generation of appropriate collection requirements. Planning activities are initiated also to acquire and organize the resources (people, skills, equipment, space, etc.) needed to implement effectively the analysis techniques that were chosen. The techniques are then implemented in the design substage.[6] The choice

----

[6] The amount of time spent in the Analysis phase will vary dramatically, of course, depending on whether the problem is well-defined and simple, so that it can be solved directly by applying a well-known algorithm [illegible] (at the other extreme) the problem is complex and ill-de-

INTRODUCTION

FIGURE 1 -- PROBLEM-SOLVING MODEL FOR INTELLIGENCE ANALYSIS



substage deals with collating and organizing the information (results and conclusions) that were generated, and using them to identify implications of potential importance to consumers.

The intelligence substage of the Presenting Conclusions phase is then initiated to chose the most appropriate format(s) for communicating the interpreted conclusions to consumers and to plan and coordinate the support resources that will be needed to produce the chosen communication vehicles. The design substage involves the actual preparation of the finished intelligence products (be they publications, cables, memos, briefings, video disks, etc.). The choice substage is then entered and the conclusions and implications are packaged in the manner chosen and conveyed (as an integral part of this substage) to the consumer(s).

If any of the substages are not completed--because new information redirected the problem-solving activity or because an insurmountable roadblock was encountered--the process recycles to an earlier substage, or, if more appropriate, the problem-solving process is terminated completely.

The treatment presented here does not pretend to supply all of the appropriate do's and dont's for the problem-solving activities suggested. Primarily it can only recommend a general attitude of actively making sure that relevant consumers get and understand the significance of appropriate information, and continually endeavoring to find ways of improving the quality and effectiveness of finished products.

THE PLANNING STAGE

OVERVIEW

The throughness with which activities during this stage are completed can impact substantially the extent to which subsequent problem-solving activities will be successful. Often it is appropriate to spend as much as one third of the total time available for an analytical project on the activities associated with this phase.

If an insignificant problem, or one that cannot be solved, is chosen for analysis, considerable time and effort can be wasted in trying to solve it. Similarly, resources will be wasted if poor communication or lax definitional activities result in the wrong problem (e.g., the wrong aspect

----

fined, so that elaborate and creative problem-solving techniques must be applied just to determine *if* an appropriate solution to the problem can of an otherwise relevant problem area) becomes the focus of analytical effort. Resources also can be wasted if they are applied to a problem that has no chance of producing a useful product (e.g., one that can be solved, but which would, regardless of the conclusions reached, have no potential for impacting a consumer's decision making process).

The Planning phase deals with the use of techniques and guidelines that help to ensure that analytical projects are directed against relevant informational needs, have the potential for producing useful results, and are defined and structured to facilitate effective use of available data. That is, this phase involves recognizing and defining relevant intelligence problems, and, once identified, determining if and how the problems so identified can be addressed meaningfully. Several distinct, but interwoven, activities comprise the Planning phase. Most of them are addressed in more detail in the Planning section of this handbook. The following describes each of the Planning activities briefly, and discusses their interrelationships.

IDENTIFYING THE PROBLEM

The first activity associated with problem solving involves realizing that a problem situation exists. From the intelligence analysis perspective, this equates roughly to recognizing that a policy-relevant information gap exists. Often this realization is "imposed" by tasking directly from policy makers or upper-level managers. But not infrequently, the analyst is the one to identify an issue that needs to be addressed/analyzed.

The initial identification may start as merely an uncomfortable feeling that an uncovered or hidden message lies among the data received or anticipated. At the other extreme, it may "appear" as a formally tasked and complete problem statement. In either case, or in the more frequent, middle-of-the-road cases, one must at some point reach, either implicitly or explicitly, the conclusion that a problem situation exists. The argument here is that one should take the time to make this decision explicitly. The following checklist of questions provides a guide to determining if a "legitimate" intelligence project has been detected:

1. Does an information gap exist?
2. Does the gap need to be closed?
3. Is it potentially feasible to close the gap?
4. Is the problem within your sphere of influence?

DETERMINING INFORMATION AVAILABILITY

Once an initial problem situation (and/or problem definition) has been identified, a natural reaction is to begin gathering information relevant to the problem area. Fine and good. Most analysts do this instinctively. The argument here is that the information collection activity has a purpose distinct from that we often associate with it. At this point in the problem-solving process we are not interested in collecting information for the purpose of solving the problem. If this unexpectedly happens, that is great, but the primary concern at this point should be to ascertain the amount and the quality of information available. This will impact substantially the subsequent problem-solving strategies that can be adopted, and (depending on the time and resources available) on the completeness and scope of the analytical conclusions that can be anticipated. The information that is consolidated will be used to support all of the Planning phase activities that are discussed next.

REDEFINING THE PROBLEM

Regardless of where along the continuum of problem clarity one's initial perception of a problem-solving task occurs, it is wise, even if not obviously necessary, to make certain at this early stage of the problem-solving process, that one is about to undertake the analytical task that best utilizes existing resources. Although this may seem either trivial or unnecessary, it seems more than worth the effort to help ensure that subsequent finished products do not "fall flat" because they either addressed the wrong aspect of an otherwise important topic, or because the problem-solving activities got sidetracked onto a technical detail of a problem and lost perspective on the consumer's needs and interests.

Problem redefinition activities involve restating a problem in ways that might trigger alternative problem definitions from the one initially identified. The intent is not simply to restate the same definition in multiple ways, but rather to broaden one's perspective of a problem in order to suggest new aspects of the problem situation that, upon being uncovered, seem more important/relevant than the initially conceived one.

The techniques used for this are divergent in nature, and come primarily from the literature on creative problem solving techniques. They consist of more than classical brainstorming sessions, and can be accomplished individually or in group sessions. Depending on the complexity and degree of understanding of the problem, they can completed within a few minutes, or can occupy several days of concentrated effort. Usually, for divergent processes such as these, redefinitional activities that extend for more than an hour or so should be broken up into multiple sessions that last no more than 45 minutes to 1 hour each. After this length of active participation, creativity (divergent thinking or idea generation) usually declines in usefulness.

ANALYZING THE PROBLEM

Similar to problem-structuring, problem-analytic techniques are designed to help one "get one's hands around" a large, complex problem, and to gain a holistic perspective on the problem's component parts, their interrelationships, and the information requirements that must be met to optimally address a given problem statement. Problem analysis techniques exceed purely structuring activities in that they attempt to make explicit problem attributes that can be used to identify Analysis-Phase techniques that are appropriate to addressing the problem that has been identified. That is, they are intended to provide the user with an increased understanding of the problem as it is currently stated. This information can also be used to support the redefinitional activities just discussed, and the constraint analysis to be discussed next.

ANALYZING PROBLEM CONSTRAINTS

As one begins generating possible problem definitions, analyzing the probable scope of the problem, and identifying potentially useful techniques, a feeling is gained for the resources that will be required to undertake a complete problem solving effort. At this point it is wise to begin evaluating those factors that will (or could) constrain analytical efforts. Typically important constraints are time, resources, and information.

Potential constraints identified (e.g., a short time deadline) can impact substantially on choices made as to what problem definition should (can) be addressed, what techniques can be used, etc. For instance, if simulation modeling (which often takes a minimum of several man-months of resource investment) seems to be the best approach to analyzing a particularly complex problem situation, but only two days are available for providing a "finished" response, then clearly a faster technique must be employed. Selecting an alternate technique may require a more limited scope to be placed on the problem definition as well.

Time probably is the major constraining factor for many intelligence analysts. Both the absolute amount of time available and the adeptness with which the problem solver(s) can effectively manage (utilize) the time that is available are important. Resource constraints are often major factors also. These concern both internal resources such as secretarial support, computer time, cooperation and/or mutual collaboration of other analysts, managerial approval, availability of any specialized training that might be needed, etc., and external resources such as external research funds, cooperation of domestic experts (industrial or academic) for obtaining specialized substantive knowledge, availability of collection assets, etc. Information constraints also are a major player, and their existence is why intelligence analysts are hired. It is the job of the analyst to make the best judgments possible in situations where data are incomplete, and to effectively guide collection programs in obtaining the most useful information possible given limited time, resources, and access. The quality and amount of information available impact substantially the number and type of methodological strategies that can be adopted profitably.

The purpose of a constraint analysis is simply to ensure that one has considered, before much time and effort has been invested, all of the identifiable factors that could prevent an analytical project from being completed adequately. Thus, the constraint analysis is in a way a control valve; one that can turn a contemplated project off, or substantially redirect it. Simultaneously, it helps pinpoint those analytical "management" tasks that are potentially critical to successful project completion, and thereby alerts one to aspects the analysis that should be given special attention.

IDENTIFYING APPROPRIATE ANALYSIS-PHASE METHODS

As is discussed in considerable detail in the chapter on Selecting the Appropriate Technique(s), the initial identification of potential techniques is driven or guided by the problem analysis activities. The problem analysis helps identify the type of problem one is addressing (e.g., moderately stochastic), which immediately suggests the generic types of techniques (e.g., probabilistic modeling) that tend to match the technique type that is associated with each problem definition currently under consideration.

It is not immediately necessary to make a final technique selection. (In fact, this probably would be a bad tactic, since it could result in this choice of technique driving the selection of the problem definition.) But, the set of potential techniques can be pruned somewhat at this stage as the constraint analysis makes it clear that some techniques are not viable options (because there is insufficient time to implement them, or because adequately trained personnel are not available, etc.).

SELECTING THE PROBLEM DEFINITION

This stage uses the information generated by the earlier stages and, in concert with the analyst's knowledge and beliefs about true consumer needs, to select the most meaningful (most useful or cost-effective) problem statement to address. Consumer needs must, of course, rank high in this selection process, but realistic assessments of what actually can be achieved must be integrated into the decision process also. One alternative to keep in mind (if it is a "politically" acceptable one) is to drop the problem entirely. That is, if ones analysis indicates clearly that a meaningful product cannot result, then it may be wise to drop the problem and use available resources on other topics that can support meaningful analytical activity.

Obviously, the set of potential techniques identified in the previous step can impact the selection of the final problem definition. If one wants to address an extremely complex problem that has a high degree of stochasticity, one needs to have available techniques of comparable sophistication and appropriate purpose. If no such technique remains on the shopping list (because of removal due to constraint analyses) then the problem definition should be modified to a scope or complexity that available techniques can handle.

The point is that techniques should influence the final selection of a problem definition only to the extent of reminding the problem solver of the limitations on what can reasonably be accomplished given the constraints bounding the problem solving activities. One should not modify a problem definition simply to be able to use a favorite technique.

SELECTING ANALYSIS METHOD(S)

As the final step in the Planning Phase, the techniques (or set of methods) that seem best to match problem-solving needs and capabilities should be selected. Since the problem analysis activities will have identified appropriate techniques and the constraint analysis will have purged the ones that were not viable to consider, more detailed knowledge about the implementation of the remaining technique choices can be brought into the selection process at this point. Other considerations, such as personal preferences for or familiarities with specific (and appropriate) techniques, as well as cost benefit assessments in terms of time utilization can be entertained also. But, the primary driving factor should be the selection of an analytical strategy (specifically selected, modified, and integrated methods) that maximizes the utility obtained from subsequent analytical efforts. This is discussed in some detail in the chapter on Problem Analysis and Technique Selection, pages 50-64.

ANALYSIS PHASE

OVERVIEW

Although more activities are involved in the Planning phase, the Analysis phase probably consumes more actual analytical time. Probably 40% to 50% of the total time available typically should be spent in this phase. In practice most analysts spend far more time than this. This probably results from two main causes:

1. A tendency to neglect (usually due to an absence of training to do otherwise) the activities that should be accomplished in the Planning and the Presenting Conclusions phases, and

2. Unexpected difficulties encountered and/or a loss of proper problem perspective during the Analysis phase (often due to a failure to complete the Planning phase activities adequately).

The reason that this phase does require more time than the other phases is that this is when all of the number/concept crunching and the information sorting and integrating is accomplished. Depending on the techniques that have been chosen for use, this phase may actually require little conceptual activity (in comparison to the Planning phase). For instance, if optimization techniques are being used, the major work is simply looking up or calculating the most appropriate input parameters for the (usually computerized) optimization algorithm to manipulate. If, on the other hand, the chosen problem is at the other end of the complexity/uncertainty scale, then a high level of cognitive activity may fill the entire Analysis phase.

ASSESSING INFORMATION NEEDS

One of the first tasks on the Analysis-phase agenda is the determination of information needs (information gaps) associated with the chosen problem definition and analysis strategies. Inherent in this task is the assumption that the analytic strategy should drive the information requirements, rather than (as seems more frequently to be the case) that the collected information should drive both the analytical tactics and the conclusions. If left to their natural preferences, most humans tend to react to (make conclusions about) the information pieces that happen to be provided, rather than to identify the specific things they want to reach conclusions about, and then drive information collection activities on the basis of the data needed to make these, and only these, conclusions. This latter, "top-down" approach to information collection has the distinct advantage of preventing one from becoming lost among the barrage of often unrelated data that can accumulate. Instead, the problem (and the collection requirements) have been structured to facilitate an appropriately integrated and orchestrated use of the data that are received. It also facilitates keeping track of just what information is still missing, and the value of that information to the quality of the finished analytical product.

Several techniques are useful for assessing information value and needs. A few are included in the Planning chapter that covers Problem Analysis and Redefinition Techniques. See, in particularly, the section on Information Value Assessment on page 71.

COLLECTING INFORMATION

Information need assessments directly facilitate the development of prioritized collection guidance packages. These can make collection activities more cost effective, as well as improve the analytical management and evaluation of information that is received. At this point, therefore, analytical collection activities become linked closely to implementation of the analytical strategy that has been selected. Each item of relevant information received should, if the problem has been structured effectively, have a predetermined place in the overall analytic picture of the problem situation and should have a relatively definable impact on current "best-guess" assessments of specific portions of the methodological "model" (whether mathematically or subjectively) formulated by the analyst.

PREPARING METHODOLOGICAL RESOURCES

This step is in many ways a managerial (project officer) function. It is focused on identifying anticipated resource needs, gaining access to those resources not directly under your own control, and attempting to manage these resources so that they are available and effective when and as needed.

A frequently useful technique for supporting this task is the Program Evaluation Review Technique (PERT). It is covered as a sequencing technique starting on page 82.

IMPLEMENTING THE ANALYSIS STRATEGY

Depending on whether one has chosen an algorithmic recipe with which total familiarity exists already, or instead has committed to using a new and sophisticated technique(s) that will require specialized training prior to full implementation, this Analysis-phase activity can vary from duck soup to severe cognitive trauma. Regardless of which extreme is closer in a given case, the steps to be taken at this stage should be clearly understood by this point in the problem-solving sequence. The steps may be quite difficult and time consuming, but they should be clear. The major question during this activity should be the extent to which collection activities can close the information gaps that have been specified. It is this uncertainty which leads us into the final activity of the Analysis Phase.

PROBLEM SOLVING TECHNIQUES
The Problem-Solving Process

SUMMARIZING AND INTERPRETING RESULTS

Two major tasks should be attended to during this stage:

1. Consolidating the results of the implemented analysis techniques into an integrated and meaningful whole, AND
2. Interpreting these conclusions in terms useful to the consumer.

That is, at this point it really is not good enough simply to regurgitate the results output by an analytical model. Nor is it acceptable simply to summarize them in a few bottom-line, major findings. One should make the effort to interpret the significance of the major findings in terms of interest to the consumer. For instance, if the results of a study indicate that a missile has three stages and can carry a payload of X kilograms, etc., it will not do the policy maker nearly as much good simply repeating these results, as it will to elaborate as to what this means in terms of the missiles usefulness as a weapons system (or space launch vehicle). Even stopping at this point is less informative than extending the analysis to placing a perspective for the consumer as to what this means in terms of a change in the overall strategic (or tactical) capabilities of the country that has developed the missile.

PRESENTING CONCLUSIONS PHASE

OVERVIEW

The Presenting Conclusions phase represents explicit recognition of the principle that one can communicate more effectively if one knows what is to be said before starting to say it. Thus, this phase, and indeed, the entire problem-solving process that has been discussed, is directed toward organizing the problem-solving process so that one will know when the problem has been solved (or, if stymied, at what point) and what the appropriate conclusions and interpretations are before trying to communicate them in finished form.

Often the urge to get started writing the final product results in analysts analyzing as they write. Frequently the results are disjointed and/or major rewrites ensue. By waiting until the content of the final product has been adequately identified, one usually can save time compared to starting to write (prepare briefings, etc.) too early in the game.

This problem-solving phase always seems to take more time than it should. Actually it should be given a reasonable amount of time. Probably 20-25% of the total problem-solving effort. Typically, people tend to grossly underestimate the actual amount of time that is required to finish

IDENTIFYING PRESENTATION ALTERNATIVES

Analysis frequently is directed toward some type of programmed production. Most other analysis is short-fuse, but directed toward a specific and well defined presentation media (e.g., the NID). In either case, if one has followed the problem solving sequence to this point, the content of the finished product should be known already. This gives added flexibility compared to the analyst who tries to analyze as he writes. Namely, one can meaningfully assess whether the initially intended presentation form is actually the best one and, if it is not, choose a more appropriate one. This way, the finished product will be able to utilize the most effective communication device, rather than having the style requirements of the initially identified forum dictate the length, breath, and audience of a finished product. (This is not to suggest that we do a poor job currently of matching finished product topics to presentation forms--we know our jobs and the publication types well enough that this probably is not the case--but, in principle, one should ask this question explicitly.)

Perhaps more useful than simply questioning the continued appropriateness of the initial choice of presentation format, is attempting to identify alternative communication vehicles to supplement the primary one. These supplemental means might be briefings, phone calls, formal or informal memoranda, etc., with such intended purposes as broadening the class of recipients, providing information tailored specifically to the needs of selected consumers, or reducing the transmission time of important details.

COORDINATING SUPPORT RESOURCES

Activities for this task resemble closely those for the corresponding planning and analytical management tasks discussed for the Analysis phase. Here one is concerned with trying to orchestrate the various players needed to get the finished product out the door and to the consumer. Typical candidates include managers and staff who must review drafts, colleagues who must coordinate, editors, and graphics personnel who must prepare visual aids. The main task is simply to make sure that each player will be available when needed, and to identify ahead of time (so that possible alternatives can be approved) situations where unacceptable delays seem likely (e.g., a division chief who has an extremely full calendar for several weeks just when you expect a draft to reach him).

PREPARING PRESENTATION VEHICLES

This stage actually comprises several subtasks. The chapters on Writing and on Briefing Techniques, as well as the one on Graphical and Visual Aids discuss them. Summarized very briefly, one is, at this stage, primarily concerned with a) getting one's thoughts together, b). getting them down on "paper", and c) getting them polished. The "system" can provide a fair degree of structure and a goodly amount of support for some of the activities undertaken at this stage.

COMMUNICATING PRODUCTS TO CONSUMERS

Simply getting a finished product "blessed" for publication, or giving a scheduled analyst briefing does not really fulfill all analytical obligations. An active role should be played by the analyst in making sure that primary consumers actual receive and understand the analytical product. This means making sure that both levels of certainty and uncertainty associated with judgements made, as well as the general direction of the conclusions, has been communicated properly. Also, one should try to determine what consumer questions remain or were generated by or during the project. These latter activities suggest more than an active one-way communication process from analyst to consumer. That is, this stage also presents the opportunity for the analyst to get feedback from consumers as to how a given product was received, how useful it was, and (most importantly) how future products might be improved.

PERSONALITY TYPES

OVERVIEW

Numerous psychological tests have been developed to identify the preferred life styles and values of people. These tests (most psychologists prefer to call them instruments) tend to be descriptive, rather than prescriptive, in nature. That is, although they do "predict" the personal characteristics that tend to belong to each personality type they identify, they do not suggest that any one personality type is better (or worse) than the another. They do not identify "good" and "bad" types, but rather try to describe the traits that belong to each type and describe how these traits impact on a person's interaction with others, the external environment, and his internal environment.

The personality profile scheme described below is based on the work of Kiersey and Bates.[7] It is the same scheme developed by Myers[8] but uses a shortened test. The personality types derived by and used for each test are identical. In fact, some places give the Myers-Briggs test but use the Kiersey verbage for handouts to explain the results.

THE KIERSEY TEMPERAMENT SORTER

The Kiersey Temperament Sorter separates personality types on four preference scales. The resulting sixteen basic personality types are identified by the set of 4 letters that indicate the end of each scale that a person tends to prefer.

In addition to the personality type, a set of ratios is given also. These ratios indicate the strength of the preference indicated on each of the four preference scales. That is, the Kiersey Temperament Sorter does not suggest that a given person is all of one or the other of a given temperament pair. Rather, each person is a blend of both. The ratio that corresponds to a given temperament scale indicates what that blend is. For instance, on the S/N scale 20 points (if all of the "S/N" questions were answered) are divided between the "S" and "N" scores. An S/N ratio of 15/5 indicates a moderately strong "S" preference. A ratio of 9/11 would indicate a slight "N" preference, but suggests a personality that displays both "S" and "N" characteristics.

---

[7] Kiersey, David and Marilyn Bates, Please Understand Me: An Essay on Temperament Styles. Prometheus Nemesis Books, Del Mar, CA: 1978.

[8] Myers, Isabel. Manual: The Myers-Briggs Type Indicator. Consulting Psychologists Press, Palo Alto, CA: 1962.

If a ratios for any of the four temperament scales is one (1.0) (e.g., an E/I ratio of 5/5), then the four-letter code includes an "X" in the corresponding position (e.g., XNTJ), and the person would be predicted to exhibit traits from both preference types (e.g., both INTJ and ENTJ for someone with an XNTJ type).

THE TEMPERAMENT SCALES

THE E/I TEMPERAMENT SCALE

The EXTRAVERSION versus INTROVERSION scale needs clarification. It does NOT relate to the popular definition of extraverts as people who are the "life of the party" and introverts as "bookworms." Rather, it is based on the technical definitions of these terms which relate to how a person prefers to recover energy. Extraverts tend to refresh their spirits and draw energy by being around people. Introverts tend to prefer solitude when they feel emotionally drained or mentally tired. Prolonged interactions with people, particularly strangers, depletes the reserves of an introvert. Extraverts, on the other hand, need such socializing to keep from feeling lonely. While the extravert is social, the introvert is territorial and needs private places in the mind and in the environment.

Words and phrases that Kiersey and Bates use to distinguish this temperament type are:

EXTROVERSION	INTROVERSION
Interaction	Concentration
External	Internal
Breath	Depth
Extensive	Intensive
Multiplicity of relationships . . .	Limited relationships
Expenditure of energies	Conservation of energies
Interest in external events	Interest in internal reaction

INTRODUCTION

THE S/N TEMPERAMENT SCALE

The SENSIBLE temperament wants facts, trusts facts, and remembers facts. He believes in experience, both personal and historically recorded. He lives in the present, concerned predominantly with things as they are. The INTUITIVE person ("N") lives in anticipation. He is primarily concerned with possibilities, with what can be, rather than with what is. The intuitive prefers to live in his imagination and thus may be less aware of reality than "S"-type persons. Intuitives often think of themselves as innovative, and skip from one activity to another (often completing none). Others are left to reap the benefits of past inspirations while the intuitive is off looking for new fields to plow.

The sensible person, in contrast, is practical. When work is involved no nonsense is tolerated. Details are not overlooked and attention often is directed to specific issues or elements of a problem rather than, as an intuitive would, to generalities of the problem as a whole.

Words and phrases that contrast the S-N temperaments are:

SENSING	INTUITION
Experience	Hunches
Past	Future
Realistic	Speculative
Perspiration	Inspiration
Actual	Possible
Down-to-Earth	Head-in-the-Clouds
Utility	Fantasy
Fact	Fiction
Practicality	Ingenuity
Sensible	Imaginative

Since the S-N temperament affects the way one looks at the world, this preference, more than any of the others, impacts a person's ability to understand and communicate with others. According to Kiersey and Bates, differences in this temperament are the source of most miscommunication, misunderstanding, vilification, defamation, and denigration.

THE T/F TEMPERAMENT SCALE

THINKING types prefer impersonal, objective methods for making a choice. FEELING types prefer to make personal, value-based judgements. The more extreme feeling types are put off by rule-governed choice, whereas the strongly thinking types may regard emotion-laden decisions as muddle-headed.

PROBLEM SOLVING TECHNIQUES

Personality Types

Feeling types usually are perceived to be more emotionally sensitive than Thinking types, but this actually is not the case. Both types react with the same emoticnal intensity, but the "T" types do not tend to let their emotions show through. Consequently, thinking types often are described as cold and unemotional. The "T" person often feels embarassed by an intense show of emotion while the "F" person may enjoy the experience.

Differences in this temperament between two people, if recognized, can be complementary more so than differences in the other temperaments.

THINKING	FEELING
Objective	Subjective
Principles	Values
Policy	Social Values
Laws	Extenuating circumstances
Criteria	Intimacy
Firmness	Persuasion
Impersonal	Personal
Justice	Humane
Categories	Harmony
Standards	Good or Bad
Critique	Appreciate
Analysis	Sympathy
Allocation	Devotion

THE J/P TEMPERAMENT SCALE

JUDGING persons prefer closure. PERCEIVING types prefer to keep things fluid and to keep their options open as long as possible. Judging types tend to establish deadlines and to take them seriously. Perceiving types tend to treat deadlines as alarm clocks that can easily be reset or which signal time to get started (rather than complete) a project. When passing deadlines to subordinates from his own superiors, a P often will move the the deadlines artificially ahead (assuming that the subordinates will not take them seriously), whereas a "J" will tend to communicate the actual deadlines.

J's tend to share a work ethic that says "work comes before all else." P's tend to share more of a play ethic. J's will tend to do anything that is useful in getting a job done (preparation, maintenance, cleaning-up), whereas P's probably will balk at anything not directly necessary to get the job done. P's are much more insistent that the work process itself be enjoyable. P's are process oriented and J's tend to be out-come oriented.

Words and phases for this temperament are:

INTRODUCTION

JUDGING	PERCEIVING
Settled	Pending
Decided	Gather more data
Fixed	Flexible
Plan ahead	Adapt as you go
Run one's life	Let it happen
Closure	Open options
Decision-making	Treasure hunting
Planned	Open ended
Completed	Emergent
Decisive	Tentative
Wrap it up	Something will turn up
Urgency	There's plenty of time
Deadline!	What deadline?
Get the show on the road	Let's wait and see...

THE FOUR MAIN TEMPERAMENTS

The sixteen types (and a metaphoric label that Kiersey and Bates give as a "best-fit" occupation to approximate some of the characteristics of each type) are:

INTP (Architect)	ESFJ (Seller)
ENTP (Inventor)	ISFJ (Conservator)
INTJ (Scientist)	ESTJ (Administrator)
ENTJ (Fieldmarshall)	ISTJ (Trustee)
INFP (Questor)	ESFP (Entertainer)
ENFP (Journalist)	ISFP (Artist)
INFJ (Author)	ESTP (Promotor)
ENFJ (Pedagogue)	ISTP (Artisan)

Kiersey argues that the sixteen personality types can be understood fairly completely if divided into four main groups. The variations among the differing types within each of the main groups are, relatively speaking, only minor variations on the theme.

PROBLEM SOLVING TECHNIQUES

Personality Types

SP'S

Persons having an "S" and a "P" preference comprise about 38% of the general population. The four SP types (ESTP, ISTP, ESFP, and ISFP) have strong similarities that overshadow the differences resulting from disagreements among them on the E/I and T/F scales. These SP types are the action-oriented, do-it-now (BUT MY WAY) types of the world.

The primary concern of an SP is to do what he wants when he wants to do it. Obligations and confining circumstances are to be avoided. Why plan, save, or wait, tomorrow may never come? Enjoy today, enjoy now, and most important, ENJOY.

Duty, Power, and Spirit are secondary. Action's the thing, and the Action must be the end in itself. Who cares if it produces a useful product? Production, completion, and accomplishment are only incidental by-products of action. They are not all that important themselves. Thus, for the SP, his "work" is essential, but not the results of the work. An SP is PROCESS oriented. The work, the action, the "whatever-is-being-done" must be fun in and of itself. If it isn't, then change is in order.

In a word, SP's are IMPULSIVE. They want to be impulsive. They may even feel guilty if they don't have impulses. What's more, the SP acts on his impulses, rather than like the rest of us, restraining actions in the name of Duty, etc.

Despite the apparent lack of dependability, an SP can be a true asset in a goal oriented organization. For, if the "work" is fun and entertaining, then the SP is not only capable, but likely to "attack" the work full force for incredibly long stretches at a time. Why not? The length of the task is of no importance. The SP does not spend time worrying about whether the work will last an hour or a day or a year. If the work is fun, doing it now is all that matters. Hence, SP's often become the true greats in things that require extremes of endurance. Their spontaneous nature enables them to rebound quickly from defeats with few, if any, scares and once again be ready for action.

Do not expect SP's to get excited about complex problems of motivation. Whatever is, is. That is sufficient for the SP. Also, expect SP's make make more mistakes than the types who plan ahead. The SP lives in and concentrates on the Present. In making judgements, an SP tends to trust Luck at least as much as rational preparation.

INTRODUCTION

Personality Types

SJ'S

Those with SJ preferences comprise about 38% of the general population. These are the duty oriented people of the world. SJ's live to be useful. The SJ must belong, and this belonging has to be earned. He is not a freeloader and will strive not to be on the receiving end of charity. An SJ may feel guilty if he finds himself dependent on someone else. He wants to be the giver and the caretaker, otherwise he would be derelict in his duties.

Hence, SJ's live in a world ordered by the rules of society. It is a world of "shoulds" and "ought-to's." The SJ thus is the personal freedom opposite of the SP. The SJ is Bound and Obligated to everything around him by THE rules. The SJ serves these obligations religiously. It is the classic Work Ethic (as contrasted to the play ethic of the SP).

Also in contrast to the optimistic SP, the SJ will tend to be pessimistic. Everything must be planned for. BE PREPARED. The SJ also is the SP's opposite in terms of stability. The SJ fosters stability. The SJ's are the maintainers of institutions. Things should be done the way they are "supposed to be done." An SJ would never dream of trying to change the RULES. The rules govern his actions, he does not govern them.

Because of his concern with belonging, titles and rank are important to the SJ. These are proof of approved and officially registered acceptance and authority. Possession is not 9/10ths of the law, the LAW is 9/10ths of possession. Accordingly, "illegal" possession is a major evil to an SJ and ranks with his giving-receiving ethic in importance.

There is no limit to the responsibilities an SJ is willing to accept. If a task needs doing, an SJ feels an obligation to see that it gets done, even if he already is overworked and doing more than his fair share. Despite this willingness to perform, the SJ frequently gets less than his deserved level of appreciation. This may be the result of the serious, often stern appearance, that SJ's seem to maintain. Their deep concern for others and for the institutions that maintain them are often hidden by this outward demeanor.

NT'S

Persons with an NT temperament comprise only about 12% of the general population. Their major function in life is Understanding. They are fascinated with Power. Not power over people, but rather power over Nature that results from being able to understand, control, predict, or explain realities. Since these are the major goals of science (control, understanding, prediction, and explanation), the NT is the ultimate stereotype of a scientist.

PROBLEM SOLVING TECHNIQUES

Personality Types

The NT loves intelligence, and defines this to mean doing things well under varying circumstances. The extreme NT can get "hooked" on the very process of acquiring intelligence (knowledge). The NT has a compulsive drive, an obsessive need to be competent. In direct contrast to the SP, who uses his abilities as a means to achieve his main objective of Action, the NT is the opposite. To the NT, action (performance) is simply the means to achieve ability. It is the acquisition of abilities to do things that is of utmost importance to the NT.

Because of his constant attention on his own abilities, an NT tends to be highly self-critical. Tell an NT that he lacks spontaneity or responsibility and he is likely to agree that you have a good point there (and not be offended in the slightest by your comment). But tell an NT that he is incompetent or stupid, and you'll quickly find out how little license he gives you to make those judgements. Only he can judge his own abilities accurately, and this he does with ruthless intensity.

Instead of the "shoulds" and "oughts" of the SJ, the NT's life is full of "should know's" and "should be-able-to's." The NT tends to accumulate massive lists of these items, and is inclined to add items without ever deleting any. He runs a "bureaucracy of excellence," and can become compulsive and a perfectionist if placed under too much stress.

Rank and formal authority have little importance to the NT. A person's worth is determined solely by his competence. The NT will judge the ideas and dictates of others himself. The social station of the source has no bearing on the soundness of an idea or opinion.

NT's a are serious appearing lot. Knowing logically that recreation is necessary for health, the NT carefully schedules his "play," and then proceeds to use this "relaxation" period to make sure that his recreational skills are improved. Errors of logic and execution are not to be tolerated (in himself) even when "playing."

The NT can appear arrogant, making others feel as if they are the NT's intellectual inferiors, even while they are being expected to live up to the perfectionistic goals of the NT (which the NT himself cannot, of course, achieve). The result of this is that others can, over time, withdraw (both socially and intellectually) from the extreme NT. This effectively isolates the NT from the intellectual experiences of others, and causes him to form inaccurate opinions of others based on incomplete and trivial data.

The NT rarely wastes words. His communications are terse, compact, and logical. Redundancy is absent or rare. The NT tends never to state the obvious; his listeners would be bored with this, of course. His choice of language is precise, and he places little reliance on non-verbal signals. The NT often is oblivious to "emotional" messages in other's communications.

More than any of the other types, NT's live in their work. Work is work, but play is work too. The worst punishment would be a condemnation to idleness. The NT must be doing something to improve. And this compulsion to improve extends to the NT's environment. Wherever they are, whatever they do, an NT will try to reorganize and improve it. Whereas the SJ's are the perpetuators of societies and institutions, NT's are their designers, creators, and perturbors. The NT listens attentively to new ideas and accepts change without stress (as long as it makes sense).

Having modified his environment or created a new institution, the NT has little inclination to stay around and run it. His concentration is on the future. He is interested in possibilities, not in current realities. Having mastered or completed one task the NT is ready to move on to the next challenge. In fact, the NT often becomes so involved with trying to organize the information he receives into a logical format that can be used to create the future, that he completely misses (or forgets to live) the events going on around him.

NF'S

The NF personality differs completely from the other three types. Where the SPs, SJs, and NTs pursue ordinary goals, the NF's goal is not at all straight-forward. The NT wants more than anything else to become "the real me." He consequently spends his life on stage, watching himself perform and trying to decide if that is the "the real me" or just an act. The NF wants to achieve self-actualization; but his focus is on the process of "becoming" that self-actualized person.

The NF spends his life seeking an identity. He wants his life to be meaningful, not just to be a small variation on a theme similar to the millions of others around him. He wants his life to make a contribution. This drive causes many NFs to enter media related fields, where their impact on others can be dirproportionately large compared to the roughly 12% of the general population that they comprise.

Along with an NF's drive to make a contribution goes his tendency to get caught up in a cause. If the cause does not provide a deep and lasting significance, however, the dedication may be shortlived. As soon as the disenchantment happens with one cause, the NF is out searching for another, more profound applications through which he can achieve self-actualization and express his unique identity.

The significance ethic of the NF demands that the work in which he engages be significant also. It is not enough that the work provide a service (which would suffice for an SJ). The work also must make a real contribution, and (in contrast to the NT temperament) the work must be important to others as well as to himself. Because of this the NF often has difficulty placing limits on the amount of time and energy he devotes to his work. Unlike the SP who can work on impulse, the NF must work toward a goal of perfection. Even though the finished products can never live up to the magnificence of its conception, the NF cannot treat his efforts as a mere commitment to complete a product. Once involved, the NF can be unwilling to limit his commitment, and can make unreasonable demands of both himself and of those around him.

Despite his tendency to attack his work with extreme passion, the NF can be an intellectual butterfly, flitting from idea to idea. Unlike the NT, who demands understanding and competency in each and every intellectual pursuit, the NF wants to taste abundantly from all of the fruits in the garden. He wants these samplings to be filled with meaning, but is not concerned about the conceptual understanding of all his intellectual contacts, as long as he has experienced them and they have provided meaning.

Like the NT the NF is focused on the future. But instead of having the NT's focus on the possibilities of principles, the NF focuses on the possibilities of people. He is more interested in people watching than in abstractions, and enjoys trying to "bring out the best" in others.

HOW TO TAKE THE KIERSEY TEMPERAMENT SORTER

The Kiersey test is available in written form in the Kiersey and Bates book referenced earlier. The Analyst Training Branch in OTE also has copies of the test.

VM users can take the test on-line by linking to and accessing the OTR 211 disk

LINK OTR 211 211 RR

ACC 211 C/A

After accessing the 211 disk, simply type in the word KIERSEY. The test will provide interactive instructions after this.

It takes about 25-30 minutes to complete the KIERSEY TEMPERAMENT SORTER.

PROBLEM IDENTIFICATION AND REDEFINITION TECHNIQUES

OVERVIEW

Redefinition techniques are designed to help clarify or expand upon the perspective currently held of a problem that has just been perceived, or that has been tentatively identified. The objective of using these techniques is to help ensure that the "correct," or at least an appropriate problem, is identified and evaluated. The techniques presented here largely assist in creating a shopping list of alternative problem statements; they are less helpful in providing guidance as to which of the generated alternatives is the most useful to select. (Problem analysis techniques, such as the one offered in the Problem Analysis and Technique Selection chapter, starting on page 50 of this Handbook, are directed more specifically toward technique selections.) Redefinitional techniques usually do, however, make explicit many of the problem attributes that need to be considered when selecting the final problem statement.

Redefinitional techniques are used heavily in creative problem solving situations, and therefore are developed most fully in the creative problem solving literature. The redefinitional techniques presented here are selected from those suggested by VanGundy*.

For many problems, the distinction between redefining and analyzing will be somewhat blurred. The act of analyzing a problem often will lead to a new definition, and redefining a problem often will involve analysis of major problem elements. The objective of both processes is to develop a new way of looking at a problem. The principles that underlie both methods involve achieving a new perspective on a problem. The basic principle underlying the redefinitional techniques is one of developing a degree of remoteness from the original problem statement. By escaping from initial problem perceptions, the problem solver is forced to use divergent thinking. Yhis should lead to more unique problem solutions. Problem analysis techniques, on the other hand, have as their underlying principle the factoring of a problem into its major dimensions or elements; thereby helping to structure or organize the information available about a problem as well as to suggest new information that might not have been previously considered. Both processes are important, but neither one is capable of guaranteeing that a "correct" problem definition will be produced. What should be produced is a better understanding of what the problem involves and a new way for viewing it.

Problem analysis and redefinition (along with technique selection) probably are the most important activities of the problem-solving process. Since their products largely determine the nature and quality of solutions

---

* VanGundy, A.B. Techniques of Structured Problem Solving. New York: Van Nostrand Reinhold Company, 1981.

that follow, any new problem statement should be considered tentative. Next to failing to adequately analyze a problem, the most common (and perhaps costly) error made by many problem solvers is rigidly sticking with the first new viewpoint that emerges. This error is especially significant for indeterminate or ill-structured problems, where there can be considerable change over the problem solving process in the amount and quality of information available.

BOUNDARY EXAMINATIONS

To define a problem one must make certain assumptions about problem boundaries. That is, one must decide (implicitly or explicitly) what portions of the environment will be considered (and what portions will be excluded from consideration) when addressing a problem. The "boundaries" that delimit that portion of the problem environment judged to be problem-relevant determine how the information you use will be organized, how it will be processed, and, eventually, how the problem itself will be solved. Proper problem boundaries also can make it easier (or harder/impossible) to solve a problem. If the boundaries are initially assumed to be closed and unchangeable, then it is not likely that they will be modified during subsequent problem solving activity. If defined too narrowly, necessary information or analytical options and solutions can be overlooked. Although it is difficult to determine the correctness of chosen boundaries, there is no question that boundary rigidity can affect problem-solving success adversely--especially when ill-structured problems are involved. Therefore, initial analyses of problem boundaries should be considered tentative. New information could always arise that would require a restatement of a problem's boundaries.

According to de Bono[10], boundary examinations do not examine the assumptions underlying the boundaries that have been chosen. Rather, the objective of boundary examinations is restructuring the assumptions of a problem to provide a new way for looking at it. The basic steps are:

1. Write down an initial statement of the problem.
2. Underline key words and phrases and examine them for any hidden assumptions.
3. Without considering the validity of these assumptions, identify any important implications they suggest.
4. Write down any new problem definitions suggested by the implications.

As an example of using this technique, consider the following problem statement: Determine the potential impact of Japanese recombinant DNA research on the United States. To begin, we write down this problem statement with the key words and phrases underlined.

---

[10] de Bono, E. Lateral Thinking: Creativity Step by Step. New York: Harper and Row, 1970.

PLANNING ACTIVITIES
Problem Identification and Redefinition

Determine the potential impact of Japanese recombinant DNA research on the United States. We now examine each of the underlined items separately to see if any new thoughts are triggered. (Frequently this will be the case, since considering each item separately tends to take it out of the context of the problem and hence helps to free one from the hidden assumptions imposed by one's initial perception of the problem's boundaries.) For instance, consider the term research in the above problem statement. Does this mean basic research only, applied research, or both? Are industrial applications included or excluded from research. When considering Japanese, does this imply interest only in research conducted on Japanese territory, or instead does it refer to work conducted by Japanese researchers, or does Japanese sponsored research conducted abroad count too, even if it is sponsored by a predominantly Japanese, but officially multinational company home-based in Paris. At this point, we can see that several hidden assumptions were present in the initial problem statement. One was that the term research was sufficiently descriptive to define clearly which activities were, or were not to be included in the study. A second, similar assumption was made for the use of the term Japanese. If the problem statement was self-generated, it dealer's choice as to which definition is chosen for each item (although one may need to make certain that such decisions about problem redefinitions are consistent with the understanding of managers who must approve programmed production topics). Hopefully, this exercise will have accomplished more than simply a clearer problem statement for others to read. It also should have triggered in your mind possible aspects of the problem that might otherwise have been overlooked. For example, one may not have considered the possibility of including Japanese sponsored research conducted in foreign locals until this redefinitional technique caused the question to be asked. One may choose to not include such research in the ensuing analytical project, and thereby end up with essentially the initially envisioned project; but the point is that the final problem definition will have been arrived at through conscious choice.

If the problem statement was generated by a consumer, it might be wise to determine what the actual scope of the problem should have regarding the terms "research" and "Japanese." It's still dealer's choice in many regards; but this time the consumer was the dealer. Problem redefinition activities may help the consumer consider aspects of a problem that were not evident before, and certainly should help clarify communications regarding expectations about the scope of the proposed analytical study.

The major strengths of Boundary Examinations are their potential for: (1) producing more useful or appropriate problem definitions, (2) clarifying often indistinguishable problem boundaries, (3) demonstrating the importance of formulating flexible problem definitions, and (4) coping with tasking that is overly precise in its problem definition. The major weakness of the technique is its lack of structure for specifying how boundary assumptions should be analyzed. In this regard, the technique is probably more a way of thinking about a problem than it is a step-by-step redefinitional procedure.

PROBLEM SOLVING TECHNIQUES
Problem Identification and Redefinition

GOAL ORIENTATION

Goal Orientation is a way of thinking about a problem for the purpose of clarifying its goals or objectives. Like most problem solving techniques, it was developed for managerial style problem solving, where the eventual goal of the problem solving effort is to take action to "correct" the perceived problem situation. For analytical purposes, goal orientation can be used in either of two different ways: 1) in its normal mode to help clarify and/or identify appropriate goals for the problem-solving effort, or 2) in an analytical mode to assess the actual goals of a foreign decision maker or decision-making group. It is implemented by considering the needs, obstacles, and constraints of a problem. The following steps are used:

1. Write down a general description of the problem. (Try to include all pertinent information.)
2. Ask: What needs to be accomplished (is needed)? What barriers might prevent satisfactory completion of the problem-solving task (obstacles)? What limitations are there on problem-solving options (constraints)?
3. Using the answers to the above questions as guidelines, write down possible redefinitions that bring the scope of the problem-statement more in line with the most appropriate, but still feasible goal.

As an example of this technique, consider the following problem statement and list of pertinent information:

Assess whether the Soviets will invade Poland?

Pertinent Information (Assume August 1981 timeframe):
The Polish economy is stuggling and there has been much internal labor and political unrest. Although the socialist political system in Poland has adopted several "liberalizing" policies, such as secret balloting, there have been no significant indications that any break with the USSR is contemplated. The Soviets are concerned with the political and military strength of the Poland, and maintaining the integrity and loyalty of the East-Block countries. Soviet troops have increased in numbers along the Polish border and among the support or advisory troops stationed within Poland. Soviet and East-Block financial support of Poland has caused limited economic strain for several Warsaw Pact countries.

Let's first examine this problem from the normal goal orientation perspective of the problem solver. To do this we examine the needs, obstacles, and constraints associated with the problem.

Needs:
Upon reflection on perceived consumer needs, we may decide that the information needed has to do with probable Soviet actions (or non-actions) to intervene in internal Polish political and economic affairs.

PLANNING ACTIVITIES
Problem Identification and Redefinition

Obstacles:
Major obstacles to completing this analytic study meaningfully may be assessed to include poor access to information about the perceptions of the Soviet Central Committee members about the severity of the Polish situation, and uncertain estimates as to the actual number and readiness status of the Soviet troops located in the vicinity of Poland.

Constraints:
A major constraint on undertaking a comprehensive study of the Polish situation may be the determination that the only analyst throughly familiar with the Polish economy has just left for a three-week trip.

Based upon these needs, obstacles, and constraints, the problem might be redefined as:

(1) Determine the Soviet Central Committee's assessment of the consequences of a political or economic collapse in Poland.
(2) Determine the capability of the Soviet military to accomplish and maintain an occupation of Polish territory.
(3) Assess the impact that Soviet military action against Poland would have on other East-Block countries.
(4) Assess the current and longer-term strength of the Polish economy under varying assumptions about sources and amounts of economic aid or loans.
(5) Assess the reaction of the Polish people to highly visible Soviet intervention in Polish internal political or economic affairs.

By redefining the problem in this way, new strategies might be suggested for solving the problem. Thus, one solution might be to ignore the actual economic situation in Poland and focus entirely on the Soviet Central Committee's perceptions of the situation and (based on their perceptions) the probable reactions that the Soviets might take.

If we use goal orientation in the second mode suggested above (in a role-playing mode) then we would attempt to make a goal orientation assessment of the problem from the perspective of the Soviets. Thus, the needs, obstacles, constraints examination might look like:

Needs:
To maintain Poland as a viable military ally and geographical buffer against foreign invasion.

Obstacles:
Liberally oriented labor movement in Poland that has been strenghtened by poor economic conditions.

Constraints:
Strained Soviet and East-Block economies that cannot make large financial or food commitments to Poland without imposing domestic hardships in these areas.

PROBLEM SOLVING TECHNIQUES
Problem Identification and Redefinition

Based on these assessments of needs, obstacles, and constraints, we would complete this role-playing exercise by listing possible problem definitions that the Soviets might select. These hopefully will help to identify aspects of the problem situation that are novel, or at least different, from those initially envisioned.

According to Rickards[11] Goal Orientation is more an attitude than it is a technique. In general, the process of applying the technique should help to create the attitude needed. Perhaps the most important consideration is to keep an open mind about exactly what the problem is and how it should be solved. Otherwise, the natural tendency will be to redefine and solve the problem using the preconceived ideas. The technique should be useful for forcing consideration of major problem needs, obstacles, and constraints. The most fruitful application probably is as a problem-broadening device at the outset of the problem-analysis stage.

REVERSALS

Every well-defined problem has an opposite and equally well-defined flip-side. In many situations a reversed definition of a problem may be more useful than the initially conceived definition. For instance, reversing the direction of a problem definition might lead to an innovative analytical approach, or might serve simply to remind one of certain aspects of a problem that might otherwise get overlooked or shortchanged because perspective had been lost on their proper relationship to the entire problem situation. For example, if one is attempting to assess the capabilities of a newly developed foreign weapons system, it would be easy to forget to address the systems vulnerabilities and design weaknesses along with the strengths that may have been a primary concern initially. Similarly, if one is considering a project to assess the possible impact that OPEC decisions could have on the US, it might be useful to consider an oppositely directed problem statement such as: What actions by non-OPEC countries will influence OPEC decisions.

Although this is an extremely simple technique, it is to this writer one of the most useful of the redefinitional techniques available. This judgement is made partly because the technique is easy to learn and to use, but more importantly, also because it is extremely effective at accomplishing the primary objective of redefinition activities. It changes dramatically one's perspective of a problem situation, often simultaneously changing perspectives about problem boundaries, analytical goals and consumer needs. It is difficult to switch thinking from positives to negatives, from strengths to weaknesses, etc. without gaining new insights or new perspectives about a problem situation. This is precisely what problem reversal techniques require one to do.

---

11 Rickards, T. Problem-Solving Through Creative Analysis Essex, U.K.: Gower Press, 1974.

PLANNING ACTIVITIES
Problem Identification and Redefinition

VanGundy[12] lists the following basic steps for using the Reversals technique:

1. State the problem as originally defined.
2. Reverse the problem statement in any way possible. (The type of reversal is not as important as rearranging the information about the problem situation.)
3. State the new definition of the problem and examine its practical implications.
4. If practical implications are not evident, reverse the problem in a different way.
5. Continue reversing the problem until a satisfactory definition is produced or it becomes apparent that no new insights will be achieved via this technique.

HYPOTHETICAL SITUATIONS

John Arnold[13] developed an effective way of redefining problems (especially technically oriented ones). His method involves constructing a hypothetical situation using the question: What would happen if...? Arnold developed a mythical planet named Arcturus IV and detailed its particular characteristics to a group of engineering students. Arcturus had such properties as volatile atmospheric conditions, a gravitational pull eleven times that of the earth's, and bird-like creatures with three fingers on each hand who were the planet's inhabitants. Arnold instructed his students to design cars, appliances, and other devices that the bird-like creatures could use. The intent of such an exercise was to remove a problem from its normal environment and create a new problem perspective that could lead to more novel problem solutions.

There are three basic steps involved in using Hypothetical Situations:

1. State the problem.
2. Develop a hypothetical situation related to the problem; the situation should not be too remote from the problem, and it should be within the problem solver's area of competence and availability of required resources.
3. Using the new problem constraints created by the hypothetical situation redefine the problem.

As for many redefinitional techniques, the basic mechanism of this technique is the programmed use of fantasy. The primary effect of fantasy is to remove the problem solver from the immediate situation and force consideration of the new problem constraints. The problem solver then can

---

12 VanGundy, A.B. Techniques of Structured Problem Solving. New York: Van Nostrand Reinhold Company, 1981.

13 Arnold, J.E. The Creative Engineer. Yale Scientific Magazine. March, 1956. pp. 12-23.

use these new constraints to consider alternative problem definitions. The Hypothetical Situation approach differs slightly from many fantasy oriented techniques, in that it uses a situation that has some direct relation to the general problem area. As a result, there is a more immediate force fit between fantasy and reality. In contrast, other fantasy methods use a more gradual and indirect forcing together of fantasy and reality. Such a difference in process does not necessarily make one approach better than another, since both have advantages. The indirect approach is more likely to produce highly unique problem perspectives while the more direct approach will make it easier to apply the perspective to reality.

When using Hypothetical Situations, the important considerations are to construct situations that are: (1) directly related to the problem, (2) realistic in terms of known laws of nature, and (3) suitable for the problem solver's level of expertise and available technological resources.

ANALOGIES

An analogy is a statement about how objects, persons, or situations are similar in process or relationship to one another. Analogies are reflected in statements of comparison such as: This organization operates like the military. Although the similarities probably hold only for specific aspects of the comparison (e.g., for speed of movement, but not for shape or for size) they have usefulness for redefining problems or generating ideas about problem approaches. Their usefulness stems from their ability to create novel "angles of view" of a problem situation by moving the problem solver to a new, but similar vantage point. This movement--a feature gives analogies a "life" of their own.[14] Analogies involving the gills of a fish or the porosity of a material could be used to provide a fresh look at illegal emigration or escape across international boundaries. By repeatedly relating such analogies to the problem and gradually developing the process or relationships, a new problem definition or solution could emerge.

The steps for using Analogies are:

1. State the problem.
2. Think of an object, person, or situation and relate it to the problem in the form of an analogy.
3. Progressively develop the analogy, translating it back to the original problem at each stage of development.
4. Continue developing the analogy until a satisfactory definition of the problem is achieved.

----

[14] de Bono, E. Lateral Thinking: Creativity Step by Step. New York: Harper and Row, 1970.

It is not necessary to select analogies that are very similar to the problem. Analogies that are remote from a problem are more likely to produce new ways of looking at a problem. Furthermore, analogies that possess many different functions, processes, and relationships also are not necessary since these can be developed while relating the analogies to the problem. The analogies should be selected on the basis of their concreteness, familiarity to the problem solver, and possession of some type of movement.

A major problem in using the analogies is finding something from which to develop an analogous relationship. The search for analogies is, in itself, a problem-solving activity. Perhaps the most important consideration is to select analogies that have a definite "life" of their own. The movement generated by such analogies is much more likely to lead to a new problem definition. As a general rule, new sources of analogies should be sought for each problem.

METAPHORS

A metaphor is a figure of speech applied to something that is not literally applicable. To say that someone has "a will of iron," for example, is a metaphorical statement since the words are used to provide a special effect in describing someone's strength of character. Metaphors are very similar to Analogies and, like Analogies, can be used to create a fantasy situation for gaining a new perspective on a problem.

Jensen[15] has described five different categories of metaphors (see Table 3) that can be used to help redefine problem statements: metaphors of (1) restoration, (2) journeys, (3) unification, (4) creation, and (5) nature. Other metaphors can, of course, be used also. Jensen's categories are representative examples that provide an initial shopping list. To use Jensen's categories:

1. Write down a general statement of the problem.
2. Select a category presented in Table 3 and examine the examples given for each subcategory.
3. Using these examples, try to develop a new definition of the problem.
4. Repeat Steps 2 and 3.
5. Examine the new problem definitions and select the one(s) most likely to produce the best problem definition.

The relationships between the USSR and each of the East Bloc countries might be metaphorically compared as a Unification metaphor, with each of the satellites cast as offspring to the single parent USSR. Viewed from this perspective, one might think of the intra-Warsaw Pact power struggles

---

15 Jensen, J.V. "Metaphorical Constructs for the Problem-Solving Process." Journal of Creative Behavior, 9: 113-123, 1975.

as typical sibling rivalries and the normal desire of growing youngsters to acquire more independence (without loss of the support benefits that have previously been provided by the parent). Viewed from this perspective, conflicts with Western countries might be envisioned as causing family crises that pull the family more closely together. Whereas isolated successes by a given country might be viewed as increasing sibling rivalries and jealousies, and/or parental resentment, thereby weakening the family unity. Cast in this light, one might decide to metaphorically "study" the dynamics of the Warsaw Pact relationships from the perspective of of a psychologist observing the interpersonal relations and group behaviors of a single-parent family.

TABLE 3 -- CATEGORIES OF METAPHORS

Major Category	Subcategory	Examples
Restoration	Medical	Ills, Rash, Cancer
	Theft	Stolen, Robbed
	Repairman	Repair, Broken down
	Cleansing	Clean up, Dirty
Journey	On Land	Barriers, Maze, Clogged channels
Unification	Family	Home, Offspring
	Shepard	Flock
	Sports	Team, Out of bounds
Creational	Edifice	Foundation, Planks
	Weaver	Fabric, Weaving
	Music Composer	Disharmony, Symphony
Nature	Light-Dark	Sheds light, Eclipse, Dull Bright
	Physical phenomena	Whirlwind, Oasis
	Biology	Cobra, Monster

Like analogies, metaphors are useful for developing flexible problem definitions that can produce unusual ideas. The development of a practical metaphor, however, is often a tricky task that requires some creative problem solving in itself. The fantasy element needed for metaphoric thinking is foreign to many persons. It may require some practice to master.

CONSTRAINT ANALYSES

TIME CONSTRAINTS

GENERAL

Time constraints are among the most critical constraints for intelligence analysis. Sufficient time to collect all data needed to make an absolutely certain judgment rarely is available. This forces intelligence judgments into the world of uncertainty, and perhaps more than any other factor separates intelligence assessments from academic studies (for which publication times are determined largely by the time at which an acceptable degree of knowledge about the subject mattex has been achieved, rather than by deadlines externally imposed by the needs of policy makers. Thus, it can be argued that time constraints force intelligence analysis (if it is to remain useful to real-world decision makers) into the world of educated guessing, and out of the realm of cautious and methodical scientific study.

Although fairly straightforward to understand and plan for, time constraints often are difficult to assess accurately. If only a few hours or a day is available to prepare an answer, clearly there is a limit on the amount of problem-solving activity that can be undertaken. Intuitive guessing could be used, but there certainly is insufficient time to build, debug, and exercise a sophisticated simulation computer model. If months or years are available to address a problem, accessing numerous data bases, tasking collection systems, and contracting for external expertise become feasible.

The point is that these tradeoff questions can (and should) be made consciously. This is nothing more than time management for analysis activities -- optimizing time utilization, or maximizing the amount accomplished per unit of time available. In addition to the information provided here, some of the techniques covered elsewhere in this handbook can be useful time management tools also (e.g., see the section on Information Value Assessment -- page 71, and the section on Program Evaluation Review Technique -- page 82).

PROBLEM SOLVING TECHNIQUES

Constraint Analysis

TIME MANAGEMENT

Methods for Saving Time[16]

1. Skim books quickly; look for ideas. (Technical or scientific literature, etc. are exceptions.)

2. Use "waiting time" to plan, to do something, or to relax.

3. Keep your watch three to five minutes fast.

4. Keep your life-goals and your personal goals up-to-date. Revise frequently.

5. Keep blank index cards with you to record ideas and to make notes.

6. Try to avoid procrastinating. If you recognize this as a pattern, ask yourself "Why?" Then confront the issue(s).

7. Begin the most important items first. Often the others may become unnecessary.

8. Arrange your "best" hours of the day to do you most challenging or creative project.

9. Delegate as much as is reasonable; but remember that delegating uses up someone else's time. Consider whose time it is.

10. Make use of the assistance of staff and other specialists to assist with special problems.

11. Make a list of specific things to do each day. Arrange in priority order. Give priority to completing the most important each day.

12. In managing your "in-basket" try and handle each piece of paper one time only.

13. Establish dead-lines for yourself and others.

14. Reward yourself on accomplishment.

15. Remember the 80/20 rule. (80% of the value can be found in 20% of the material.)

---

[16] Adapted from a list compiled by Alan Lakein, Professional Time

PLANNING ACTIVITIES

Constraint Analysis

Identifying Time Wasters

Below are listed the time wasters most commonly encountered in consulting on time management with senior executives.[17] To assist the reader in analyzing his own time wasters, possible causes and solutions are suggested for each. These are not intended to be exhaustive but merely to serve as guidelines for further diagnosis. Causes and solutions tend to be personal, while the time wasters themselves are universal in nature.

Time Waster	Possible Causes	Solutions
Lack of planning	Failure to see the benefit	Recognize that planning takes time but saves time in the end.
	Action orientation	Emphasize results, not activity.
	Success without it	Recognize that success is often in spite of, not because of, methods.
Lack of priorities	Lack of goals and objectives	Write down goals and objectives. Discuss priorities with subordinates.
Overcommitment	Broad interests	Say no.
	Confused priorities	Put first things first.
	Failure to set priorities	Develop a personal philcsophy of time. Relate priorities to a schedule of events.
Management by crisis	Lack of planning	Apply the same solutions as for lack of planning.
	Unrealistic time estimates	Allow more time. Allow for interruptions.
	Problem orientation	Be opportunity-oriented.

---

17 MacKenzie, R. Alec, The Time Trap, Managing Your Way Out, AMACOM, New York, 1972, pp. 172-176. The list is adapted from Managing Time at the Top (New York: The Presidents Association, 1970).

PROBLEM SOLVING TECHNIQUES

Constraint Analysis

	Reluctance of subordinates to break bad news	Encourage fast transmission of information as essential for timely corrective action.
Haste	Impatience with detail	Take time to get it right. Save the time of doing it over.
	Responding to the urgent	Distinguish between the urgent and the important.
	Attempting too much in too little time	Attempt less. Delegate more.
Paperwork and reading	Knowledge explosion	Read selectively. Learn speed reading.
	Computeritis	Manage computer data by exception.
	Failure to screen	Remember the Pareto principle. Delegate reading to subordinates.
Routine and Trivia	Lack of priorities	Set and concentrate on goals. Delegate or drop nonessentials.
	Oversurveillance of subordinates	Delegate; then give subordinates their head. Look to results, not details or methods.
	Refusal to delegate; feeling of greater security dealing with operating detail	Recognize that without delegation it is impossible to get anything done through others.
Visitors	Enjoyment of socializing	Do it elsewhere. Meet visitors outside. Suggest lunch if necessary. Hold stand-up conferences.
	Inability to say no	Screen. Say no. Be unavailable. Modify the open-door policy.
Telephone	Lack of self-discipline	Screen and group calls. Be brief.
	Desire to be informed and involved	Stay uninvolved with all by essentials. Manage by exception.

PLANNING ACTIVITIES

Constraint Analysis

Meetings	Fear of responsibility for decisions	Make decisions without meetings.
	Indecision	Make decisions even when some facts are missing.
	Overcommunication	Discourage unnecessary meetings. Convene only those needed.
	Poor leadership	Use agendas. Stick to the subject. Prepare concise minutes as soon as possible.
Indecision	Lack of confidence in the facts.	Improve risks as validating procedures.
	Insistence on all the facts--paralysis of analysis	Accept risks as inevitable. Decide without all facts.
	Fear of the consequences of a mistake	Delegate the right to be wrong. Use mistakes as a learning process.
	Lack of a rational decision-making process	Get facts, set goals, investigate alternatives and negative consequences, make the decision, and implement it.
Lack of delegation	Fear of subordinates' inadequacy	Train. Allow mistakes. Replace if necessary.
	Fear of subordinates' competence	Delegate fully. Give credit. Insure corporate growth to maintain challenge.
	Work overload on subordinates	Balance the workload. Staff up. Reorder priorities.

DAILY WORKSHEET/TIME SCHEDULE

ITEMS TO DO		EARLY MORNING
		8:00 8:15 8:30 8:45
		9:00 9:15 9:30 9:45
		10:00 10:15 10:30 10:45
		11:00 11:15 11:30 11:45
	DELEGATE TO	12:00
MUST DO		12:15 12:30 12:45
		13:00 13:15 13:30 13:45
		14:00 14:15 14:30 14:45
SHOULD DO		15:00 15:15 15:30 15:45
		16:00 16:15 16:30 16:45 17:00
		LATE AFTERNOON

PROBLEM SOLVING TECHNIQUES

Constraint Analysis

RESOURCE CONSTRAINTS

Resource constraints concern the amount and quality of support that is available to accomplish a task. Do you have to work alone, or can other people (library personnel or other reference personnel be used? If they can be used, are they available? What is the quality and quantity of the work produced by others and by yourself? Does the project require computer, word processor, or other equipment support? If there is time, are funds available to provide external contract support?

When selecting the methodological strategy to be followed, one needs to assess if personnel having specialized methodological skills are needed. If needed, are they available or can other, available personnel be trained in time? Depending on the answers to these questions, adjust your (and your consumer's) expectations as to the scope, timeliness, and quality of the finished product.

A simple guideline exists for evaluating resource availability: If the resources (equipment, money, personnel, etc.) required to solve the problem are not available and are not likely to be available in time to solve the problem, terminate the process. If the process cannot be terminated (i.e., an answer must be provided) then the process must be adapted. Simpler, probably less effective, but also less resource intensive, strategies must be adopted. A critical factor at this point is being able to predict if resources not needed at time X and not available then are likely to be available at time Y when they are needed. Of course, lack of such information would in itself indicate an information-resource deficiency. Furthermore, the seriousness of this deficiency must be weighted against the importance of solving the problem. If the problem is seen as important and uncertainty about the availability of resources can be tolerated, then a decision could be made to continue the problem-solving process.

INFORMATION CONSTRAINTS

DATA AVAILABILITY

The amount and kind of information available can significantly impact the completeness and the acceptability of any problem-solving effort. Insufficient data can prevent a meaningful assessment from even being attempted (although the consumer may still require a response, even if it is only your' best guess). Too much information can degrade results also, because one can, if not careful, get lost in the mass of detail and lose sight of the often broader-scoped consumer needs.

To this end, it is wise to evaluate the relative worth of differing kinds of information (e.g., is something truly problem related; and if numerical data, on what kind of measurement scale was it based). By making these kinds of evaluations, one can help to restrict your data to useful categories, and can better determine what strategies or methods would be most appropriate for analyzing it.

TYPES OF DATA

Nominal Scale

A nominal scale separates data into distinquishable categories. The categories have no attributes which provide a natural order or mathematical hierarchy by which the observations can be ranked. (For instance, blood-type, sex, religion, hair-color, and citizenship are nominal categories.)

Ordinal Scale

An ordinal scale separates data into categories that have a natural order, but insufficient specificity to permit comparisons of the relative size of differences between the various categories. Cold-hot-warm and small-medium-large represent ordinal classification schemes. (Note that it is not clear whether the difference between cold and warm is the same as, or is greater or less than the difference between warm and hot.)

Interval Scale

Interval scales place observations into categories that have both natural orders and defined numerical relationships. For instance, temperatures measured on the Centigrade scale indicate not only which of two temperatures is higher, but also how large the difference is.

Ratio Scale

Ratio scales are interval scales that have an absolute zero. For example, temperatures measured on the Kelvin scale permit one to compare the ratio of the absolute values of two temperatures, as well as determining the magnitude of the difference between them. (i.e., 100 degrees Kelvin is half as warm as 200 degrees, as well as being 100 degrees cooler. This is not true of temperatures measured on the Centigrade scale.)

PROBLEM SOLVING TECHNIQUES

Constraint Analysis

"BUREAUCRATIC" CONSTRAINTS

Bureaucratic constraints perhaps apply most to self-initiated projects. Will management support and/or accept the project? Do the relative efforts involved in editorial process between two differing publication types make one format more desirable than another? But even projects tasked from above may have hidden bureaucratic constraints if, for instance, it becomes obvious that some plausible results of your analysis would be unacceptable.

A basic question here is whether the problem is within your "sphere of influence." This addresses the question of whether actions can be taken without receiving authorization from superiors or others (e.g., peers, subordinates, experts) who might need to be involved. If the problem is not within this sphere of influence, superiors, or relevant others must be consulted, their authorization or approval obtained, and their appropriate involvement determined. Should consultations with others reveal that a problem is totally outside of a problem solver's sphere of influence, the process would terminate unless approval to proceed is provided. Thus, a general guideline for this decision point would be: Consult with others whenever their authorization or approval is needed to solve a problem. If required authorizations or approvals are obtained, assume that the problem is within your sphere of influence; otherwise, terminate the problem-solving process. If there is no need to consult with or rely upon others and all of the preceding constraints have been satisfied, then proceed on the assumption that a potentially solvable problem exists.

PROBLEM ANALYSIS AND TECHNIQUE SELECTION

APPROACHING ANALYSIS HOLISTICALLY

The concepts of interdisciplinary research and interdisciplinary analysis have attracted increased interest over the past decade. But the anticipated benefits, expected to arise through the simultaneous and coordinated use of the diverse methodological tools of differing disciplines, have emerged only stubbornly. One reason appears to be the absence of a key ingredient needed for success in interdisciplinary endeavors: the absence of a holistic, or system's oriented, methodological approach.

Few individuals, however skilled they may be in their own specialties, have received training in or become skilled at selecting or developing syncretic methodological approaches that are congruent to an overall problem situation. Analysts usually specialize by expanding knowledge about substantive parts of a problem at the expense of gaining a comprehensive perspective on how the myriad parts fit together to operate as a whole. Those missing, syncretic skills, are relatively straightforward to acquire and moderately easy to learn. The following highlights a general systems theory approach to adopting an effective, eclectic problem-solving strategy that is syncretic. It is based heavily on the detailed text on this subject by professor John W. Sutherland.[18]

In developing a general systems theory (or perspective) for intelligence analysis, we shall adopt the three "dictates" that Sutherland ascribes to general systems theory:

1. The proper focus is holistic.

2. The selection of analytical tools should be eclectic, based solely on the needs and properties of the problem situation.

3. Common isomorphisms among problem situations should be exploited.

In following these three dictates, we shall develop what is primarily a generic, problem-analysis methodology. It is based on the first two dictates, which in turn supply the philosophical base from which the third can be attained. The goal will be to bring increased structure to the problem-solving process, without bringing so much that real complexities are sacrificed to analytical expediency. To achieve this goal an abstract, but unified, conceptual framework for identifying the analytical requirements and limitations of various types and forms of systems will be developed.

----

18 Sutherland, John W. Systems: Analysis, Administration and Architecture Van Nostrand Reinhold Company, New York, 1975.

PROBLEM SOLVING TECHNIQUES

Problem Analysis and Technique Selection

The development of this analytic system approach will start with a discussion of the structural and behavioral dimensions that can be used to classify problem-types. The discussion will then extend into a description of several typical (Ideal-Type) problem situations and examine them for those specific classes of analytic methods that are appropriate matches. Finally, the issue of selecting the appropriate technique(s) for a given problem situation will be addressed. This primarily involves attempting to identify the most appropriate technique that can be brought to bear on a problem given the constraints (lack of time, limited resources, absence of personnel trained in desirable techniques) that restrict the set of techniques that can realistically be implemented.

MEASURING PROBLEM COMPLEXITY

The basic strategy for identifying problem-appropriate techniques involves analyzing a problem situation to determine general attributes that make some classes of techniques more relevant than others. One begins by performing this analysis on the problem situation as initially understood, but (as problem redefinition, problem structuring, and constraint analysis methods are applied during the problem-solving process) may end up either updating the analysis on an altered problem definition, or may even end up using the technique "in reverse" to determine what, if any, specific problem statement should be addressed <u>if the constraint analysis indicates that insufficient time, resources, etc. exist to solve the problem that had been selected</u>. In any case, the identification of an "optimal" technique(s) for the problem at hand relies primarily on classifying the problem situation into a general typology of problem types that have associated with them a prescriptive set of congruent analytical techniques.

According to Sutherland, problem situations can be classified into methodologically prescriptive categories based on three measures of structural and behavioral complexity. The dimensions measured are those of the problem-situation's internal structure (domain), its interactions with external factors--those lying outside the problem boundaries--that can impact internal factors (ecological), and the behavior--stable versus highly variable--of the problem situation over time (dynamic).

From an analytic perspective, a problem's degree of complexity primarily refers to its degree of analytical tractability. That is, if one can measure current problem parameters, can determine the causal interrelationships, and can predict accurately future behavior, a problem is tractable. A problem becomes less tractable, more complex, as it becomes increasingly difficult to determine the current value of internal factors (state variables), as the internal causal relationships become less well understood, as the importance of external factors becomes greater or their predictability decreases, and the expected response of internal factors to specific circumstances becomes decreasingly certain. Thus, one can sense intuitively that problem situations will be inherently simple when internal factors are relatively insensitive to or isolated from exogenous (external) factors (or when the external factors are reliably predictable or, better yet, unchanging), when the structure of the internal factors is homogeneous and symmetrical in nature, and when problem variables change little, or at least predictably, over time.

Of course, few problems lie at either extreme of the "complexity" scale on any, much less all, of the problem dimensions (domain, ecological, and dynamic). Instead, complexity on each dimension normally will lie at some intermediate level on a continuum of possible levels. Furthermore, the three dimensions chosen for classification purposes in this problem typology clearly are not sufficiently robust to detail all analytically relevant aspects of a problem situation. The reduction of analytically relevant attributes to a three dimensional representation constitutes both a descriptive simplification (any problem can be described as a set of X,Y,Z coordinates) and an assessment complication (each dimension represents a composite cr average description of multiple attributes that must be integrated equitably). Significantly, these two "approximations" do not prevent the development of an analytically useful typology.

The dimensional attributes that can be used to assess the degree of complexity associated with a given problem are summarized in Table 4 and are discussed below.

ATTRIBUTES OF THE DOMAIN DIMENSION

The domain, or structure of relationships among factors internal to defined problem boundaries, can be viewed in several generic ways to help in the assessment of complexity along this dimension. For one, the factors can be examined to see if they are discrete or continuous-state variables. Discrete variables can take on only a specific set of clearly defined values, whereas continuous variables can take any of an infinite set of values (although the range of a continuous variable may be no greater than a discrete one). For example, temperature is a continuous-state variable since it can take any value (measured to whatever number of decimal places you chose--at least in theory). Measurements of IQ are discrete, since only integral values can be obtained. Since discrete variables encompass a smaller set of possible values, they are, in principle, somewhat less complex to deal with than are continuously distributed variables. (Continuous variables can, of course, be partitioned into "discrete" groups, which allows them to be treated effectively as discretely distributed variables; but this involves making a simplifying assumption that lessens a true representation of reality and that does not need to be made when dealing with truly discrete data.)

A second structural element of the domain dimension has to do with whether the variables interact with one another in a linear or non-linear manner. This is a mathematical concept that describes whether variables that interact with one another in a cause-effect manner change according to a relationship that maintains a constant ratio (i.e., a given change in variable "A" will always cause proportionately matched percent change in variable "Y") or whether the percent change in the effected variable has

TABLE 4 -- SUMMARY OF PROBLEM DIMENSION ATTRIBUTES

DIMENSIONAL ATTRIBUTES	DEGREE OF COMPLEXITY	
	Tractable	Intractable
ECOLOGICAL		
Source of control over internal variables	Endogenous	Exogenous
Method by which external interfaces are established	Selective	Universalistic
Behavioral complexity of the system's environment	Well-behaved or Placid	Turbulent or Protean
DOMAIN		
Fraction of the domain that is empirically accessable	High	Low
Degree of structural homogeneity	High	Low
Degree of symmetry in the distribution of structural elements	High	Low
Size (Scope) of the domain	Small	Large
Density of domain elements	Low	High
DYNAMIC		
Range of behaviors available to system components	Few	Many
Number of interfaces affecting system components	Few	Many
Nature of relationship between parts and whole	Dependent	Semiautonomous
Range of response "postures" that the system can adopt to handle external inputs or output demands	One	Many

some more complicated mathematical relationship. The first case, where a constant proportion relates the percent changes in the two variables (or their inverses) is referred to as a linear relationship, since if the change in the first variable is graphed versus the change in the second variable, the points comprising the resulting scatter diagram will form a straight line. Furthermore, this line can be expressed as an equation in terms of the two variables such that neither variable is raised to a power (has an exponent) other than one or zero and neither of the variables is ever multiplied or divided by the other. The corresponding graph of a non-linear relationship between two variables will not be a straight line, and the equation expressing the relationship between the variables will involve either exponents other than zero or one, or cross-products (multiplication or division) between the variables. The linear situation is easier analytically, and hence systems with linear domains tend toward the simpler end of the domain complexity scale.

Internal organization also can be approached from a holistic perspective. Namely, when considering the overall organization of the internal factors, consider whether they are comprised of similar, repeated entities, or if instead each of the factors is sufficiently different from the others to be functionally or organizationally unique. The former, simpler type of system can be referred to as a segmented system. It might be represented as a manufacturing industry comprised of multiple assembly lines, each doing the same job as the others and each having the same number and type (job category) of personnel. To understand all of the assembly line units, one needs only to study one of them. They are functionally and structurally identical (in principle). Contrast this situation with a functionally organized bureaucracy. The names of the bureaucratic components may at first glance seem the same (the organization may be comprised of directorates that are subdivided into offices which are further broken down into divisions, etc.), but the function and, probably, the composition of each component will differ substantially. One "office" may be composed of scientists, another of maintenance personnel, and the next of logistics experts. Each probably has differing numbers of personnel assigned to it. This analytically more complex structural situation is classified as a differentiated system.

To understand all of the components in a differentiated organization, one needs to study each component individually. This obviously is more difficult than treating a segmented system, but at least some overall umbrella structure and cause-effect relationships exist. Organizations such as scientific think-tanks, or grass-roots guerrilla movements may have no detectable structure (at least from a management perspective) at all, with each individual in the system simply "doing his own thing," completely independent of what any other member might be up to. Clearly, this is an even more complex domain situation. This attribute of internal structure relates to whether there is an identifiable "vertical" structure (e.g., a chain of command). If some units seem subordinate (in a cause and effect sense) to others, with yet other units subordinate to them (such that no component is directly subordinate to more than one other component and no circular pattern exists in the cause-effect flow), then the system is hierarchical in nature. This type of structure is easier to treat analytically than is a structure which lacks a strong or clear control mechanism. Systems which have such "anarchical" cause and effect structures (which might be diagrammed as a heavily interrelated network) are classified as non-hierarchical. The think-tank example cited above (with its ambiguous control relationships) tends toward this more complex, non-hierarchical domain situation.

ATTRIBUTES OF THE ECOLOGICAL DIMENSION

PROBLEM SOLVING TECHNIQUES
Problem Analysis and Technique Selection

The ecological dimension primarily is concerned with a system's interfaces with the "outside world." Of interest are such things as the extent to which the system communicates with external factors (i.e., is it insulated from external factors--a closed system--or can external factors impact easily on internal system variables, is the system's behavior determined predominantly by external influences--exogenous control--or by internal relationships--endogenous control, and does the system exist within a stable (placid) environment or is the external environment protean and turbulent.

Clearly, problems that are highly self-contained (insulated from external effects), highly self-directed (strong internal regulation and control of behavior), and resident in a placid environment (the influences from which, even if strong, are essentially unchanging over time) will be analytically more tractable than problems tending toward the other extremes of these ecological attributes.

ATTRIBUTES OF THE DYNAMIC DIMENSION

System behaviors can be categorized into deterministic and stochastic systems. Steady state systems try to maintain constancy (tend to maintain a constant output) or seek to produce a finite state output (exhibit one of a small set of possible behaviors). The steady state system is eminently predictable in that its homeostatic behavior causes the same behavior to be observed at all times (e.g., temperature regulation in healthy, warm-blooded animals). The finite-state system, on the other hand, is somewhat more complex. System behavior will alter over time; but the range of behavior changes will be restricted to a, presumably, small set of possibilities. Knowledge of the starting state and of current influences make assessment of the resulting final behavior possible.

What begins to make problem situations noticeably complicated is when system behaviors become probabilistic (stochastic). Here, the set of possible behaviors can become large and/or difficult to predict except in a statistical sense. For instance, if one rolls a single die, it is known that eventually it will come to rest with exactly one of its six sides facing up, and the the "up" side will bear a number between 1 and 6, inclusive. However, despite the fact that the number of possible outcomes is small (6), knowing the current state of the die or its previous history does not help one (assuming that it is an honest die) predict what result the next role of the die will produce. The best one can do is assign an equal likelihood (of one sixth) to each side. Thus the output is statistically predictable, but exact prediction based on cause and effect is precluded. Despite this loss in ability to predict output exactly, one is (in this moderately stochastic case) at least fairly certain that the eventual outcome has been included in the list of possible options. (It is extremely unlikely that the die will come to rest on an edge, so that no side will face up.) As the number of possible outcomes increases, the options become effectively non-finite in number since there is a corresponding degradation in ones ability to list accurately the all-inclusive set of options that will include the final output. That is, there becomes an increasingly greater chance that something not even thought of will happen. In this case, one is confronted with a situation where a large number of potential outcomes can be identified, but where little confidence exists that the list includes the scenario that actually will unfold. Problems of this type are referred to as severely stochastic.

IDENTIFYING METHODOLOGICALLY PRESCRIPTIVE PROBLEM TYPES

The three problem dimensions discussed above can be used to classify problems. One could, for example, assign an ordinal ranking from one to four to a given problem for each of the three dimensions. This would allow each problem to be classified into one of 64 (4 X 4 X 4) possible categories, each identified by one of the permutations that can be generated by the resulting X,Y,Z coordinate system. The complexity of this typology limits its usefulness.

Sutherland[19] argues that only four Ideal-Type systems are needed to capture the analytically prescriptive richness of the dimensional problem analysis that has been described. He develops these Ideal-Types in detail, and argues that they are sufficiently robust to construct the prescriptive tool we seek for selecting congruent (appropriate) methodological techniques.

The Ideal-Type systems Sutherland generates basically assume that complexity along each of the dimensions (domain, ecological, and dynamic) and along each of the dimensional attributes change in an interdependent and predictable fashion. Thus, when domain structures become more complex, so do ecological and behavioral attributes. By breaking each dimension into four "correlated" levels of complexity, Sutherland develops, describes, and then uses (in a methodologically prescriptive manner to be discussed below), the four fundamental Ideal-Types included in Table 5 and described in Table 6.

The way one determines which Ideal-Type category best matches a given problem is relatively straightforward, albeit somewhat subjective. First, examine the domain, ecological, and dynamic dimensions of a problem defined by a specific problem definition and assess the degree of analytical tractability associated with each dimension. (The dimensional attributes discussed above and summarized in Table 7 that impact the analytical tractability of each dimension can be used to aid making these assessments.) Dimensions having easy analytical tractability will be associated with deterministic Ideal-Type problems, and those tending toward analytical

---

[19] Sutherland, John W. Systems: Analysis, Administration and Archi- [illegible] New York 1975 pp. 108-142.

TABLE 5 -- SYNOPSIS OF IDEAL-TYPES

SYSTEM IDEAL-TYPE	ECOLOGICAL	DOMAIN	DYNAMIC
I Deterministic	Autarchic	Fully Segmented	Mechanical
II Moderately Stochastic	Symbiont	Structurally segmented Functionally differentiated	Nonequifinal
III Severely Stochastic	Dominant	Fully Differentiated	Equifinal
IV Indeterminate	Dependent	Structurally differentiated Functionally segmented	Heuristic

intractability, with indeterminate Ideal-Types. The Ideal-Type corresponding to the overall problem will be basically the "average" of those corresponding to the three dimensions.[20] As a final step in deciding which Ideal-Type best describes the problem one is evaluating, consult Table 6 to determine which general Ideal-Type description best fits.

SELECTING THE MOST APPROPRIATE TECHNIQUES

The rational used to match specific techniques/technique types to each of the Ideal-Type problem situations is based on the supposition that the power of a technique and the resources needed to use it should be matched to the complexity and importance of the problem at hand. That is, techniques should be congruent with the analytical requirements of the problem. Using an unnecessarily sophisticated technique results in overkill and an inefficient use of analytical resources. An error in the opposite direction (using too simplistic a technique on a complex problem) results in ineffective problem-solving since an unnecessarily high probability of error will be present.

This supposition about using congruent techniques is accepted in this text, but amplification is needed to define congruent in terms of real-world intelligence problems. This involves considering the impact that problem-solving constraints play in intelligence analysis. For instance,

----

[20] General systems theory suggests that most problems will have similar degrees of complexity on each dimension. Most problems tend to follow this rule, but even those "special cases" that do not can be treated by this Ideal-Type approach since the problem attributes that have the lowest analytical tractability will tend to dominate in selecting appropriate techniques and best-match Ideal-Types.

TABLE 6 -- PROBLEM IDEAL-TYPE DESCRIPTIONS

IDEAL-TYPE	DESCRIPTION
Deterministic	Accurate empirical data can be obtained easily, causal relationships are understood and are highly specific.
Moderately Stochastic	Empirical data is generally obtainable, but may not have the accuracy desired, some uncertainty exists about the specificity of cause and effect relationships, but events and variables (even though somewhat uncertain) will remain within prespecified and manageable range.
Severely Stochastic	Empirical measurements are difficult with variable accuracy, causal relationships are complex and understood only in part. System is quite variable and is <u>not</u> constrained to an easily manageable range of alternative responses or futures. No one obviously most likely future scenario can be identified.
Indeterminate	Data is incomplete, inaccurate, and/or inaccessible. Relationships are so poorly understood that problem boundaries are unclear. Not only are likely future alternatives unclear, but there is uncertainty as to what areas of study will prove most fruitful for reducing the uncertainty. Historical precedents (similar problem situations) are not available.

unlike the academic community, where deadlines are usually long-term and flexible, and unlike management or industrial situations, where a fair degree of problem-solver control exists over access to data, the intelligence analyst often must respond to externally imposed short deadlines on issues where little data is available or can be obtained. Thus, this text would restate the technique sophistication-problem complexity congruence supposition in a way that incorporates the (often methodologically substantial) constraints extant in the intelligence environment:

> An analytically congruent technique is one that most cost-effectively assesses a problem situation given the specific constraints impinging on the problem-solving process.

Thus, if one is required to respond analytically to a problem task that resembles a TYPE III problem Ideal-Type, but is given only four hours to provide the written product, building a stochastically based simulation model--which may be the methodologically appropriate way to go--will not be feasible. Nonetheless, one should not necessarily retreat immediately to outright intuitive guessing in such situations. For instance, adequate time probably does exist to construct a simple probabilistic model, to apply some problem structuring techniques, or--if permissible--to redefine the problem to one sufficiently simple to approach it adequately in the time available. Using this philosophical approach, one is justified in using an analytically "inappropriate" technique provided that it is the best that can be done considering the constraints impacting the problem solving effort. One is not justified in using an inappropriate technique simply because it's "easier," "familiar," "popular," "faster," etc. The method suggested here to select congruent techniques will make use of problem-specific judgments about the time, resource, and information constraints that affect analytical capabilities and specific judgments about the constraints that apply to using specific techniques.

If approached from a methodological standpoint, simple (TYPE I) problems are predominantly data driven. That is, they are amenable to empirical investigation. The data that are collected are highly accurate and the relationships that connect the problem variables are well understood. Thus, it is appropriate for one to use inductive reasoning and thereby to make generalizations about the whole from the specific information collected about the parts.

As one moves into the realm of TYPE III and TYPE IV problems, the empirical-inductive approach becomes inappropriate. The data are no longer so easily available, and the variable interrelationships are generally unknown or only understood poorly. In this situation one should adopt a hypothetical-deductive approach. This involves reasoning from the general to the specific. One starts by making hypotheses about relationships and using the best data (or guesses at data values) available, and uses this hypothetical problem system to determine the logical implications of the assumptions (hypotheses) made in constructing the "model." The analytical "accuracy" of the implications produced by the model depends on the accuracy of the hypotheses. This accuracy must be judged post-facto (and often quite subjectively) based on how well the model's "predictions" match "reasonable" real-world expectations and/or how well the model can "predict" or replicate history accurately. If the model is judged to perform well, then confidence is increased in the hypotheses made about the problem situation; but the "TRUTH" or absolute correctness of the hypotheses has not been (and via this appr ch will never be) proven conclusively.

The process of identifying analytically congruent techniques for each of the four Ideal-Types thus becomes, based on the gradual transition from inductive to deductive processes as increasingly complex problems are encountered, simply one of matching the methodologically philosophical assumptions and/or compatabilities of differing techniques with the corresponding properties of the Ideal-Types. Thus, one moves from very positivistic (operations research, optimization) techniques (where "THE" correct answer is expected) to heuristic and meta-theoretical approaches as one shifts from TYPE I to TYPE IV problems situations. Table 7 gives an overview of prescriptively appropriate techniques for each problem Ideal-

Type.[21]

By the time one has identified the Ideal-Type that most appropriately describes the problem being addressed, it is possible (using something like Table 7) to identify potentially useful techniques. To determine exactly which technique(s) should be selected requires some specific information about the techniques being considered and one's capability to implement the techniques successfully given existing problem constraints. For example, consider the following situation:

> One is confronted with a Moderately Stochastic problem of assessing the maximum production capability of some industrial plant. Adequate raw (input) materials are known to exist for any hypothesized realistic production level. The problem basically is determined to center around the through-put capacities of the multiple processes that occur (in parallel) inside the plant. These individual processes are understood well, but some uncertainty exists about the exact equipment in use at the plant in question and the manpower available for the labor-intensive (non-automated) processes.
>
> Based on this, we can assess the ecological dimension to be easily tractable (adequate raw materials, no indication--let's assume--of labor troubles or power shortages, etc.), the Domain is generally tractable but does have some uncertainty as to exactly how each of the multiple processes is conducted, and the Dynamic Dimension apparently is stable. Based on these assessments, the problem would appear to be largely Deterministic, but with some characteristics of Moderate Stochasticity.
>
> Optimization techniques (specifically network analysis and possibly queuing theory) should be appropriate. Since some uncertainty exists, a good methodological choice might be to build a simple, probabilistic simulation model. The question thus becomes one of whether this is a feasible technique to try and implement.

---

[21] Although it has not been discussed explicitly, the reader should realize that smaller "pieces" of a larger problem may belong to different (usually simpler) Ideal-Types than the larger, "big-picture problem. Thus, a Moderately Stochastic problem involving the assessment of the energy efficiency of some country might include a Deterministic subproblem about the generating capacity of a specific hydroelectric dam. This subproblem should be addressed using optimization techniques. The Moderately Stochastic categorization of the larger problem, and the association of probabilistically based techniques as congruent tools, apply to the syncretic method that should be used to pull together the diverse subproblem areas into a meaningful whole. Different teams from a common task force might thus work on their own specific problems with different "levels" of methodological tools, but the teams need to synthesize their individual results via a method that is congruent to the larger problem. It is this "umbrella," syncretic method that determines the kinds of results that will be obtained.

PROBLEM SOLVING TECHNIQUES
Problem Analysis and Technique Selection

TABLE 7 -- IDEAL-TYPE/TECHNIQUE ASSOCIATIONS

IDEAL-TYPE	TECHNIQUE CLASS	ANALYTIC BASE	EXPECTED RESULTS	ASSOCIATED TECHNIQUES
Deterministic	Optimization	Positivistic (hypothesis-free analysis)	THE ANSWER with low probability of error.	*Optimization Techniques *Normative Modeling
Moderately Stochastic	Extrapolative Projective	Inductive or Statistical Inference	Range of possible answers with some index of confidence. Chance of error determined objectively and depends on data accuracy.	*Probabilistic Reasoning *Inferential Statistics *Role Playing *Correlation, Regression, Time Series *Simple Game-Based Models *Simple Simulation Models
Severely Stochastic	Contingent/ Game Based	Deductive Inference	Array of loosely defined alternatives. High chance of error. Logical, rather than objective, probabilities are generated.	*Stochastic Simulations *DELPHI *Analogic Models *Game-Based Models *Descriptive Techniques *Multiattribute Utility Models (Role Play)
Indeterminate	Heuristic/ Metahypothetical	Intuitional/ Metaphysical	Generation of methods or heuristics to guide further studies. No attempt to be accurate or to make realistic forecasts.	*Creative Problem Solving *Heuristics *Scenario Building *Hypothetical-Deductive Analysis *Axiologic Tools

To someone with appropriate methodological skills (a knowledge of probability theory, a simulation programming language, network analysis techniques, and queuing theory) this problem would probably require one to two man-days of effort, assuming access a) to a computer that supports the simulation language and b) to the expert knowledge about the probable inner workings of the plant. If these skills and resources are available to the problem solver (either personally or among subordinates, co-workers, or support methodology groups) then this methodological approach is appropriate if at least a day or two is available to determine an answer.

If less time is available, then a "faster" (and probably less appropriate) technique should be looked for. In this case, one might decide that queuing backlogs between processes were unlikely to be a major problem, and only the throughputs of the multiple processes would have significant impact. In this case, one could use either a graphical or computerized algorithm to perform a network analysis. By redoing the analysis with "pessimistic" and "optimistic" assumptions (as well as best-guess assumptions) about the through-put capacities of the individual multiple processes, one could get both a best-guess answer to the likely plant output, as well as a general feel for the potential error in this estimate based on the uncertainties existing for each individual process. The results of this approach would be less robust than using the simple simulation model first suggested. but they still would be better than simple intuitive guessing--which always is a legitimate, last-resort fallback position when time (or another constraint) becomes too limiting.

If, in our example, adequate time and manpower exist, but methodological or substantive expertise are lacking, the options of bringing in outside help (external contractor support) or acquiring appropriate training should be explored. In terms of this example and the training option, it probably would take one to two man-weeks of training to get someone to the point that they could develop the desired simple simulation model. Attempting this through self-study probably would not lengthen the process significantly. (Similarly, if time were limited and one chose to use a graphical network analysis algorithm, this method can be learned in less than half an hour.)

The point to be made is simply that in selecting the best technique to use, a match must be made between existing resources (including man-power, equipment, knowledge, time, training alternatives, methodology support groups, external research funds, etc.) and the comparable requirements of the techniques that are being considered for use. One cannot accomplish successfully more than there are sufficient resources to support, but simplified methods usually can be developed to make maximum use of the resources that are available. The trick, if there is one, is to identify all of the resources that are available. Several methodology support groups exist within the intelligence community.[22] But even if such groups cannot provide the man-power you might desire at the time you would like it, they usually can provide guidance as to potentially useful techniques

---

[22] The CIA's Information Science Center, (703) 351-3532 has a community-wide charter.

and to sources of information about those techniques. They also should be able to assist in making judgements about the resource requirements involved with attempting to use a specific technique for a specific problem, can provide information about the kinds of results that will be produced by a given technique, and can offer guidance as to how such results should be interpreted.

IDENTIFYING APPROPRIATE RESULTS AND CONCLUSIONS

Each of the technique categories listed in Table 7 has a differing epistemological foundation. That is, each methodology is based on a specific set of fundamental assumptions about such things as how the world is organized, how much validity (firmness) should be attached to the data being used, and how reliably predictable the causal relationships specified are. Since differing methodological philosophies have been prescriptively assigned to each of the problem types, it is reasonable to expect analytical results for one problem type to differ in quality as well as in expected numerical accuracy from other types. That is, one should have differing problem-solving expectations for differing Ideal-Type problem situations. When faced with a Type III problem (and most intelligence problems fall into the Type II and Type III categories), one should not expect to achieve the same predictive accuracy that is possible with a Type I problem. Instead, one should attempt to accomplish what is appropriate to a Type III problem, and should communicate both to management and to consumers, the analytical possibilities and goals that should be associated with that particular Type III problem.

Since one can determine a problem's Ideal-Type relatively early in the Pre-Evaluation Phase of a problem-solving effort, appropriate analytical expectations can be determined and communicated sufficiently early that false expectations can be prevented or corrected before an irrecoverable and substantial analytical resource investment has been made to try to achieve results that are unrealistic to obtain. Table 8 highlights the major analytical tasks and goals associated with each of the Ideal-Types. Note that (Table 8) as one progresses from Deterministic into more and more complex (less analytically tractable) problems, the results obtained become less certain and more descriptive. With Indeterminate problems, one basically is confined to attempting to identify important factors for the consumer to keep an eye on. Trying to provide information about "THE" most likely future of an Indeterminate problem would be doing the consumer a disservice, since by definition of an Indeterminate problem, something other than what has even been thought of probably will happen. Similarly, providing a wishy-washy response to a Determinate problem fails to provide the analytic clarity that should be possible.

The assertions just made imply that the interpretations made from analytical results should be tempered according to the degree of certainty appropriately associated with those results. Thus, if the analytical results are "squishy," then several if-then types of interpretations become appropriate; and the consumer should be made aware of the fact that multiple, all reasonably feasible futures exist, and that policy actions should be taken with this factor in mind. If the problem is Deterministic, then the fact that the projected future is a relatively certain prediction should be made clearly, and the implications of "THIS" future to possible policy considerations should be assessed as completely as possible.

TABLE 8 -- ANALYTICAL TASKS AND GOALS FOR THE PROBLEM IDEAL-TYPES

IDEAL-TYPE	ANALYTIC TASK	ANALYTIC GOAL
DETERMINATE	Develop optimization algorithms or complete predictive/descriptive models.	Accurately determine specific parametric and coefficient values.
MODERATELY STOCHASTIC	Generate statistically disciplined or inductively reasoned extrapolations or forecasts.	Determine moderately tight ranges within which parametric and coefficient values should fall.
SEVERELY STOCHASTIC	Develop array of logically possible, qualitative alternative futures.	Determine the most likely and most important relations that are likely to govern future alternatives.
INDETERMINATE	Generate heuristic or other strategy to limit problem boundaries and focus future work.	Identify the set of problem variables most likely to produce analytically useful information.

ASSESSING INFORMATION NEEDS

OVERVIEW

Assessing information needs primarily involves trying to assess the relative potential worth (in terms of being able to "solve" the overall problem) of collecting differing items of information. Such assessments can be made in numerous ways, from purely intuitional guessing to building elaborate simulation computer models to perform mathematically exacting sensitivity analyses. All but the intuitional guessing approach involve placing some sort of structure on a problem. Hence, this chapter could appropriately have been titled "Problem Structuring Techniques." The title used was chosen merely to be as consistent as possible with the general problem-solving approach presented earlier in the Handbook. Note that most of the techniques presented here can become part of the overall analytical strategy used for the overall Evaluation Phase of a problem-solving effort. The techniques presented primarily are for organizing data and concepts into forms that facilitate comprehension, communication, and subsequent retrieval. The techniques covered will range from intuitive, unstructured ones to others that involve a high degree of methodological rigor. All will involve structuring, displaying, or assessing the relative value and interrelationships among differing items of information.

Information structuring techniques have much in common with display techniques when used in analysis. Although displays are used to present information to others; the process of developing each specific display involves quite a bit of structuring work. Structuring techniques, on the other hand, typically are used to organize data into an intelligible form for one's own personal consumption, often to help analyze and/or redefine an ill-structured, complex problem. The end goal of such structuring activities is thus the development of displays useful to yourself. The value assessment techniques help to identify the relative importance of differing items.

CHECKLISTS

Checklists are collections of words (but other things, such as labeled file folders or organized stacks of papers, serve similar functions) that represent ideas worth remembering. The checklist therefore serves to document these ideas in a "list" form so that they can be recalled or referred to more easily, readily, and reliably than would be the case if only personal memory were used.

Techniques for generating checklists are largely subjective. Brainstorming and creativity techniques are used frequently. To be most useful, the initial list generated via such techniques generally needs to be revised dramatically to eliminate redundancy, clarify definitions, and eliminate non-relevant and unimportant items. Often this revision process is done intuitively. During the latter parts of this section we'll suggest some alternative methods.

As an example of a simple (and for sake of space, by no means complete) checklist, consider Figure 2. You can undoubtedly add to this list. Hopefully, if you had developed it for analytical use, you would get others to add to (or delete from) it also.

FIGURE 2 -- POPULATION DYNAMICS CHECKLIST

Factors Considered Relevant to Assessing Population Dynamics

Birth Rate
Death Rate
Infant Mortality Rate
Life Expectancy
Average Family Size
Population Structure
Daily per capita Calorie Consumption
Economic Conditions
Food Import Costs
Urban/Rural Migration
Unemployment Levels
Emigration

Disadvantages associated with checklists include:

1. The checklist itself does little to stimulate its creator or user to try to identify missing items. (A task at which most people are not good.)

2. A checklist that has become "institutionally" accepted can inhibit the search for or recognition of new or emerging factors (i.e., checklists can make one less sensitive to change).

3. The technique does not itself make interactions or relationships among the listed factors readily apparent.

ANALYSIS ACTIVITIES

Assessing Information Needs

DIMENSIONAL ANALYSIS

Jensen[23] has developed a method designed to clarify and explore the dimensions and limits of a problem. The technique, Dimensional Analysis, can be used to help generate a checklist (thereby helping to overcome the third deficiency listed for checklists), as well as to decompose a problem along the various analytical dimensions it encompasses. Specifically, Jensen suggests examining problems on five dimensions:

1. Substantive dimension
2. Spatial dimension
3. Temporal dimension
4. Quantitative dimension
5. Qualitative dimension

Each of these dimensions is directed toward answering five fundamental questions: What? Where? When? How much? How serious? As shown in Table 9, the dimensions then are further analyzed by responding to a series of specific questions.

VanGundy suggests the following steps for using Dimensional Analysis:

1. State the problem.
2. Briefly write down separate descriptions of the problem in terms of What? Where? When? How much? How Serious?
3. Using these descriptions, answer the questions listed for each of the dimensions (Table 9).
4. Evaluate the answers to these questions by considering the impact that each can have on efforts to solve the problem.
5. Select those areas most important for analysis.

One difficulty in using Dimensional Analysis for analytical purposes stems from Jensen's emphasis upon using it as a management tool that focuses on human-relations problems. This focus on social and psychological problems that some selectivity be exercised in using Dimensional Analysis to analyze technical problems. The trick, of course, is knowing which dimensions and questions are relevant. Perhaps the best use of the technique would be as a checklist during pre-problem-solving activities or while selecting among alternative evaluation techniques.

<u>Substantive Dimension</u>

Commission or Omission? Does the problem being addressed involve something being done actively (e.g., that could be stopped or modified) or instead something not being done (either through neglect or intentional abstention)? Problems involving "sins" of commission are generally easier to detect and deal with than "sins" of omission.

---

[23] Jensen, J.V. "A Heuristic for the Analysis of the Nature and Extent of a Problem." Journal of Creative Behavior 12: 168-180, 1978.

PROBLEM SOLVING TECHNIQUES
Assessing Information Needs

TABLE 9 -- DIMENSIONAL ANALYSIS

Substantive	Spatial	Temporal	Quantitative	Qualitative
Commission or Omission?	Local or Distant?	Long-standing or Recent?	Singular or Multiple?	Philosophical or Surface?
Attitude or Deed?	Particular Location(s) Within a Location?	Present or Impending?	Many or Few People?	Survival or Enrichment?
Ends or Means?	Isolated or Widespread?	Constant or Ebb-and-Flow?	General or Specific?	Primary or Secondary?
Active or Passive?			Simple or Complex?	What Values are Being Violated?
Visible or Invisible?			Affluence or Scarcity?	To What Degree are the Values Being Violated?
				Proper or Improper Values?

Attitude or Deed? Sometimes problems involve attitudes in addition to, or rather than, observable events. In such situations, analyses should address policy-relevant attitudes (of foreign officials, for instance) that could be modified diplomatically, rather than concentrating solely on military, economic, or other policy leverages.

Ends or Means? Is the issue under study the result of something else, or is it the cause of a potential (or existing) problem? What is observed might be only a symptom of the "real" problem.

Active or Passive? What type of threat is posed by the problem? Some problems are threatening to the national security and well-being of the country, while other problems are just irritating obstacles that would be desirable to overcome.

Visible or Invisible? Problem situations should be analyzed to determine if the "real" problem is readily apparent or if it is hidden and likely to be overlooked. Physical problems often tend to be visible while social-political problems tend to be more invisible.

<u>Spatial Dimension</u>.

ANALYSIS ACTIVITIES

Assessing Information Needs

Local or distant? The effects of some problems are limited to a local geographical area while the effects of other problems have more far-reaching implications. If the problem is local, then the information and the implications derived from the information will be local. (i.e., what are the problem boundaries?)

Particular Location(s) Within a Location? Particularly when a problem is geographically diverse or broad, the specific locations of a problem components need to be isolated so that problem-solving activity can proceed efficiently.

Isolated or Widespread? The extent to which a problem is linked with problems in other areas must be determined to insure successful problem solving. (i.e., how interdependent versus independent are the sub-elements of a problem?)

Temporal Dimension.

Long-Standing or Recent? Some problems have been part of the world for centuries while others are a more recent phenomenon. (The "solutions" to some long-standing problems may become the causes of more recent problems.)

Present or Impending? Does the problem exist now or is it pending if present trends continue? If an existing problem, is the situation static, or could it develop into something more serious?

Constant or Ebb-and-Flow? Some types of problems occur rhythmically over time or in a specific time relation to other events. Does the problem occur on a regular (predictable) basis? Usually, the more unexpected a problem, the greater is the difficulty involved in dealing with it.

Quantitative Dimension.

Singular or Multiple? A problem can exist because of a single item or because of many items in combination. Sometimes the individual items only become a problem when considered together, and are unimportant if assessed by themselves alone. Differing items usually have differing degrees of importance to understanding the problem.

Many or Few People? This question is relative. What is small in one situation might not be so in another situation. Nevertheless, many problems cannot be adequately defined until the number of interacting persons (or number of population types or groups) has been determined.

General or Specific? How general (important?) is the problem area being addressed? Do its implications impact a general category or only specific subgroup?

Simple or Complex? Does the problem consist of only a single, isolated element, or is it made up of many, interacting elements?

PROBLEM SOLVING TECHNIQUES
Assessing Information Needs

Affluence or Scarcity? Has the problem situation been caused by an abundance or a lack of something?

Qualitative Dimension.

Philosophical or Surface? Does the problem involve philosophical assumptions?

Survival or Enrichment? A problem might be a matter of survival or merely bring into question the quality of a situation. A survival problem might be expected to result in more immediate action than an enrichment problem.

Primary or Secondary? The perceptions that the active players have about the relative importance of a problem will determine how and when the problem is dealt with.

What Values are Being Violated? While the previous questions have been concerned with obtaining facts about a problem (What is?), this question seeks to determine: What is wrong? As defined by Jensen, a problem exists whenever there is a violation of values. Jensen feels that it is important to identify the specific values being violated, and to recognize that the relative importance of values will vary from person to person.

To What Degree are the Values Being Violated? Whenever a value is violated, the implications for problem solving will be more important when the violation is serious than when it is trivial.

Proper or Improper Values? Another issue in the area of value violation is the question of whether or how a particular value should be honored. Differing players in the problem scenario may have not only differing, but even conflicting views on appropriate values.

WEIGHTED CHECKLISTS

This is less of a new technique than simply a possible variation on the theme of the basic checklist. What we suggest here is simply making greater use of visual display in writing down a checklist. That is, instead of making all of the checklist items the same size, make the items you feel to be of the most importance largest in size. The least important items should similarly be made the smallest. Alternatively, put symbols in the margin to indicate importance: e.g., 5 stars for the "biggees" to 1 or no stars for the unimportant factors.

This technique is particularly useful for displaying information to others rapidly, for it indicates both the scope of your concerns (by the items on the list) and the relative weighting you give to each of the items listed.

As an additional whistle or bell, important items (or items having a common element: economic factors versus military factors, for example) can be grouped together to make several sublists within the original checklist.

INFORMATION VALUE ASSESSMENT

The technique we shall introduce here is an extension of the basic checklist procedure discussed above. The extension is largely an intuitive one, but does add a great deal of power to the use of checklists for making cost effective use of your analytical time.

To proceed, we start with a checklist of important items and systematically evaluate at least two things about each item. One is the item's level of importance (for assessing your particular analytical problem) relative to the other factors on the list. (It usually is easiest to scale the factors on some arbitrary scale with which you are comfortable: 0-100, or 1-7, or 0-10, etc.) Start by identifying the most important factor and assign it the highest value on your scale: e.g., 100. Then compare each of the other factors to this one. If you consider them to be half as important give them half the value: e.g. 50.

The second thing to assess for each factor is the percent of unknown but obtainable information about that item that you need to know to solve your analytical problem. That is, picture in your mind a scale (such as the one in Figure 3) ranging from 0 to 100% knowledge about that item. A certain fraction of that scale represents knowledge that is known to you already. The remainder represents information that you do not have (do not know). Of this unknown fraction, a portion should be theoretically obtainable, whereas the rest is unobtainable. Of the obtainable portion, decide what fraction represents data you actually need to draw meaningful conclusions, and what fraction comprises unknown, problem-useless trivia. What you want to identify on your checklist is the fraction of the total knowledge scale that is unknown, obtainable, and is needed for successful analysis (i.e., the Unknown, Collectable, Needed Analytically fraction of Figure 3).

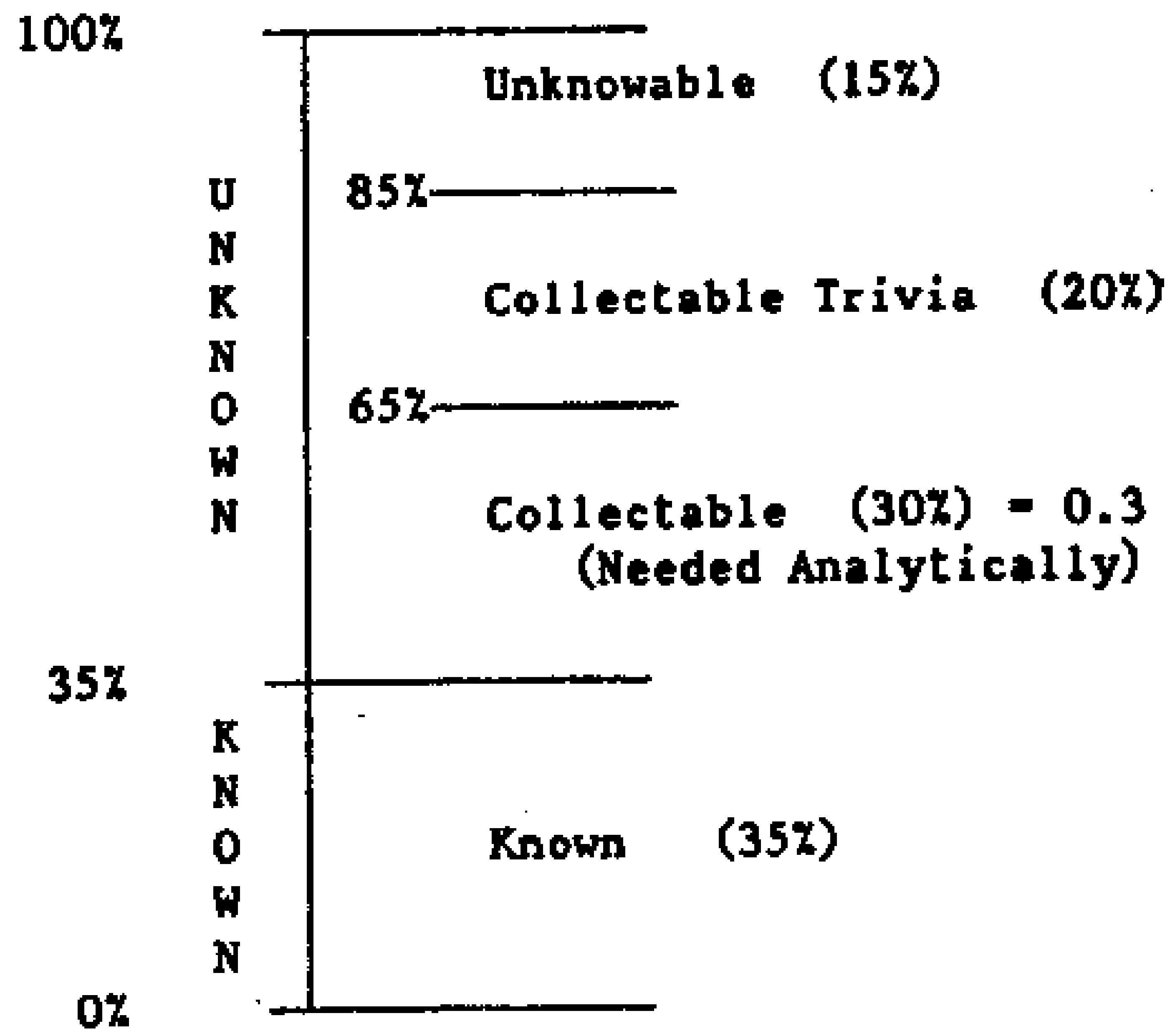
For the example shown in Figure 3 we would indicate on the checklist that 30% of the knowledge about this item was unknown but obtainable and needed. You will probably find it more convenient to enter a 0.3 than 30% since we'll be doing some multiplication later, but entering 30% or even 3/10th will work.

After completing the numerical entries in both columns by each item in the checklist we are in a position to determine the potential information gain of collecting actively against each of the items. This is done by multiplying together the two numbers assigned to each item.[24] (e.g., if the

---

[24] Implicit in this multiplication process is the assumption that the

FIGURE 3 -- COLLECTABLE INFORMATION SCALE



first item had an importance score of 30 and a score of 0.3 for the fraction of information about it that still could be obtained, then actively collecting against it would have an information value of 0.3 * 30 = 9. (See the 2nd item in Figure 4.) The 9 has meaning only in comparison to the information value scores of the other items on the list. Its relative weight should reflect the value of this item relative to the other items on the list. Remember that the "correctness" of the information value weights is only as good as the "correctness" of the numerical scores assigned earlier. This assignment process may have had verified, hard data driving it, or (more likely) have been quite subjective. Subjective assignments do not make the numbers "wrong," but do suggest that the final information value weights should be used as a guide, not as a rule book.

At the beginning of this section we raised the concept of using analytical cost effectively. Up to now the time dimension has been ignored. That is, we ignored the fact that if time is short, we may not be able to collect information needed on the more important topics.

---

problem solving value represented by the missing but collectable information is distributed uniformly (linearly) among the entire data set. That is, the first 10% of the information is assumed equally important to the last 10% (or to any given 10% out of the middle). In most real-world situations this is a weak assumption, but when this technique is used in a comparative mode, it probably is not a dysfunctional approximation (in that it most data categories will be affected similarly). If this assumption seems particularly bad for any given item, it may prove useful to further divide the item being evaluated into subcomponents [illegible] distributed value.

FIGURE 4 -- POPULATION DYNAMICS INFORMATION EVALUATION

Importance	Needed Obtainable Knowledge		Information Value
90	0.1	Birth Rate	9
30	0.3	Death Rate	9
80	0.7	Infant Mortality Rate	56
100	0.1	Life Expectancy	10
.	.	Average Family Size	.
100	0.8	Population Structure	80
.	.	Daily per capita Calorie Consumption	.
.	.	Economic Conditions	.
.	.	Food Import Costs	.
.	.	Urban/Rural Migration	.
.	.	Unemployment Levels	.
10	0.6	Emigration	6

If time is indeed a critical (limiting) factor in our analytical project, then our goal should be to obtain the most problem relevant information value that we can in the limited time we have available. The information value assessments we have made can be used to help accomplish this goal if we are willing to estimate the amount of time needed to obtain the collectable information referred to for each topic. That is, using a common measurement unit (e.g. man-days) estimate the amount of time that it will "cost" <u>you</u> to collect against each topic.[25]

For instance, assume that (referring to Figure 4) you estimate that it will "cost" you one man-hour each to determine the birth and death rates, four man-hours to determine the infant mortality rate, a two man-hours for life expectancy, ten man-hours for the population structure, and six man-hours for emigration figures. By dividing the information value figures calculated earlier by these time estimates, we can calculate the information acquisition rate for each topic. The higher this number, (shown in

---

[25] If time is a limiting factor for another player in the collection process (e.g., another analyst or a librarian) their time commitments should be estimated also in a separate list.

the last column of Figure 5) the more information value we can obtain per unit time of analytical effort.

If a project must be completed by a deadline only 8 man-hours away, and you estimate that it will take (using the numbers given in the preceding paragraph) 24 man-hours to collect everything, then the best strategy would seem to be using the eight man-hours available against the topics which offer the highest rates of information acquisition.

FIGURE 5 -- POPULATION DYNAMICS INFORMATION ACQUISITION RATES

Population Dynamics Information Evaluation

Importance	Needed Knowledge Obtainable		Information Value	Time	Info Rate
90	0.1	Birth Rate	9	1	9
30	0.3	Death Rate	9	1	9
80	0.7	Infant Mortality Rate	56	4	14
100	0.1	Life Expectancy	10	2	5
.	.	Average Family Size	.	.	.
100	0.8	Population Structure	80	10	8
.	.	Daily per capita Calorie Consumption	.	.	.
.	.	Economic Conditions	.	.	.
.	.	Food Import Costs	.	.	.
.	.	Urban/Rural Migration	.	.	.
.	.	Unemployment Levels	.	.	.
10	0.6	Emigration	6	6	1

Based on the calculations shown in Figure 5, you should first attack the Infant Mortality Rate topic, which provides 14 units of information value per man-hour. Next, the Birth and Death Rate topics should be addressed (9 units per man-hour). Thus far 6 man-hours (of the 8 available) have been used. The remaining 2 man-hours available should be used to address the Population Structure topic. This strategy would yield (according to the estimates made) a total of:

14*4 + 9*1 + 9*1 * 8*2 = 80

units of information value. No other time allocation will produce this much problem-relevant information acquisition in the eight-hour time period given.

(Though the theory sounds nice, a word of caution or perspective needs to be added here. It relates to the uniform or linear distribution of information value assumption that was discussed above. One may in fact be better off spending at least a little time on each topic about which little is known, rather than completely exhausting the information stores available on any given topic. That is, the information within each topic category may have a marginal value that diminishes as more and more is already known about the topic--the last 2 may be much less useful or needed than the first 2. Don't get trapped by the numbers calculated. They are a guide, and are not meant to replace judgment. Nonetheless, the Information Acquisition rates that were calculated should prove useful for deciding which topics to emphasize if time is limited. In the example used above, the acquisition rates calculated certainly suggest that the Infant Mortality Rate topic should be addressed much more throughly than Emigration, and that the full 8 hours should not be devoted solely to collecting against Population Structure.)

The disadvantages of this method include:

1. Interactions among items are not explicitly identified.

2. The value of information is assumed to be distributed linearly across the entire universe of problem relevant information.

3. The numerical assignment methods suggested do not take into explicit consideration the amount of information about an item that might be obtained as a predictable result of collecting against another item on the list.

PROBLEM SOLVING TECHNIQUES
Assessing Information Needs

HIERARCHICAL STRUCTURING

Assessing the importance of something, as we did with the the information value assessment method, can be done in many ways. We primarily have relied on intuitive judgment thus far. This often is sufficient for many purposes. At times, though, we may desire a more structured approach, either because the item(s) being assessed are complex in nature or because we want to make explicit how the estimates were derived.

One method to add both structure and documentation is hierarchical ordering of important aspects of the item we are assessing. This amounts to indicating, in a graphical and layered fashion, the factors which contribute to the importance of the items being assessed. The things which make each of these factors important are then listed under the factor to which they apply. The process is continued by breaking each of the current bottom "nodes" into its basic components until you are comfortable assessing directly the importance of each of the current bottom nodes. What we are building will look much like a wiring diagram for a government bureaucracy. The importance of NFAC, for instance, might be thought of as a composite of each of its offices. The importance of each office also could be represented as the resultant sum of the contribution made by each of its divisions. The process would be continued until the person (group) making the overall assessment felt comfortable assessing the value of the bottom nodes listed. This might be at the level of the divisions, the branches, individual analysts, or the offices themselves. The bottom nodes also might differ from office to office or division to division. For example, if a high level manager who had come up through the Office of Strategic Research were making the assessment, the assessments might be much more fine grain for OSR than for the other offices.

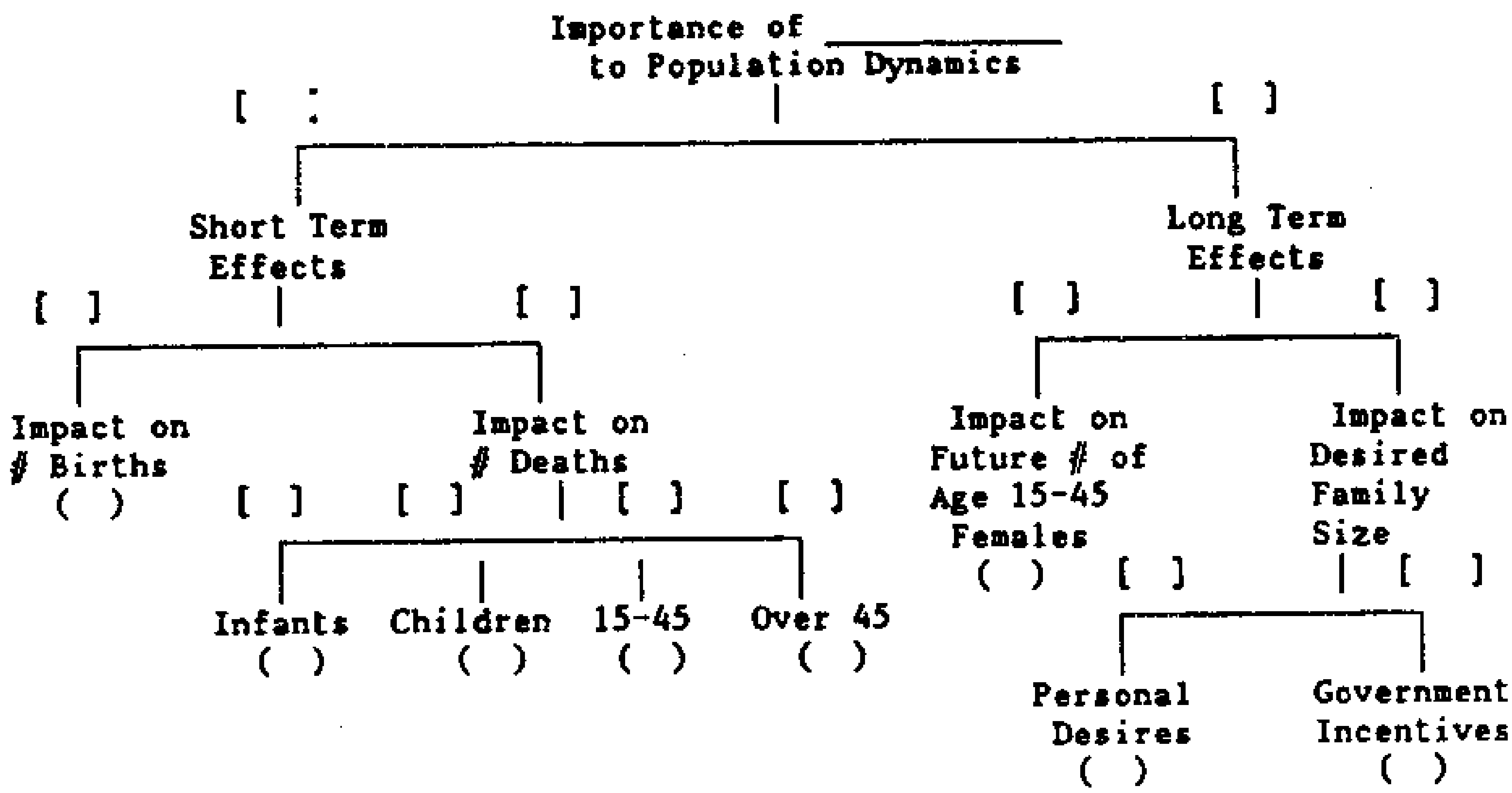
Two things must be done to reduce the hierarchical breakdown back to an overall assessment of the top node (the item we started out to assess in the first place).

1. The importance of each bottom node must be assessed. This is done by estimating (subjectively or, if relevant data exist, objectively) the impact of the potential range of variability of the top node item in terms of the bottom node factor being assessed. Assign importance values ranging from zero (0) for no meaningful impact, to one (1.0) for full impact possible.
2. The relative importance of each of the factors associated with the next higher level node need to be indicated. The sum of these "importance weights" must equal one (1) for each set of nodes linked directly to the next higher node.

As a quick example, consider the importance weight assessments we made earlier for the population dynamics value analysis. We might have decided that the importance of each item depended on too many things to make a meaningful overall assessment intuitively (the old 7 plus or minus 2 rule). In that case we might have devised something like the hierarchical structure shown in Figure 6 to help us make the overall assignment.

ANALYSIS ACTIVITIES
Assessing Information Needs

FIGURE 6 -- SAMPLE HIERARCHICAL EVALUATION STRUCTURE



Now that the structure has been determined, we can evaluate the relative importance of the factors affecting each higher level node. Remember, the sum of the factors for each node should be one (1.0). (Figure 7 gives example importance weightings.)

We now can use this structure to evaluate each of the items on our information value assessment checklist. We begin by selecting the item to be evaluated first. Then we assign importance values for this item to each of the bottom nodes. The " ( ) " indicate where these assignments are needed.

To evaluate the tree, we then multiply the values assigned to each bottom node by that node's weighting factor and (as the plus signs are intended to indicate) we add together the resultant value times weight products for each of the next higher level nodes to determine that nodes value. These calculated values are then multiplied (in turn) times their respective weights, until a final sum of products is determined for the top node. This is the top node's (item's) value.

Let's try a specific example with our Population Dynamics Tree. We'll try evaluating the importance of possible changes in Daily per capita Calorie Consumption using the assessments shown in Figure 8. Assume that we're dealing with a lesser developed country that has maintained an adequate dietary level up to now, but which could develop severe shortages in the near future. There is no chance that excess food could become available, but it is conceivable that minimal dietary levels could be maintained if everything else (weather, food prices, inflation, etc.) goes well. We need to evaluate the tree using the difference in the range of possible outcomes as our base.

FIGURE 7 -- SAMPLE HIERARCHICAL EVALUATION STRUCTURE (WITH WEIGHTING FACTORS)

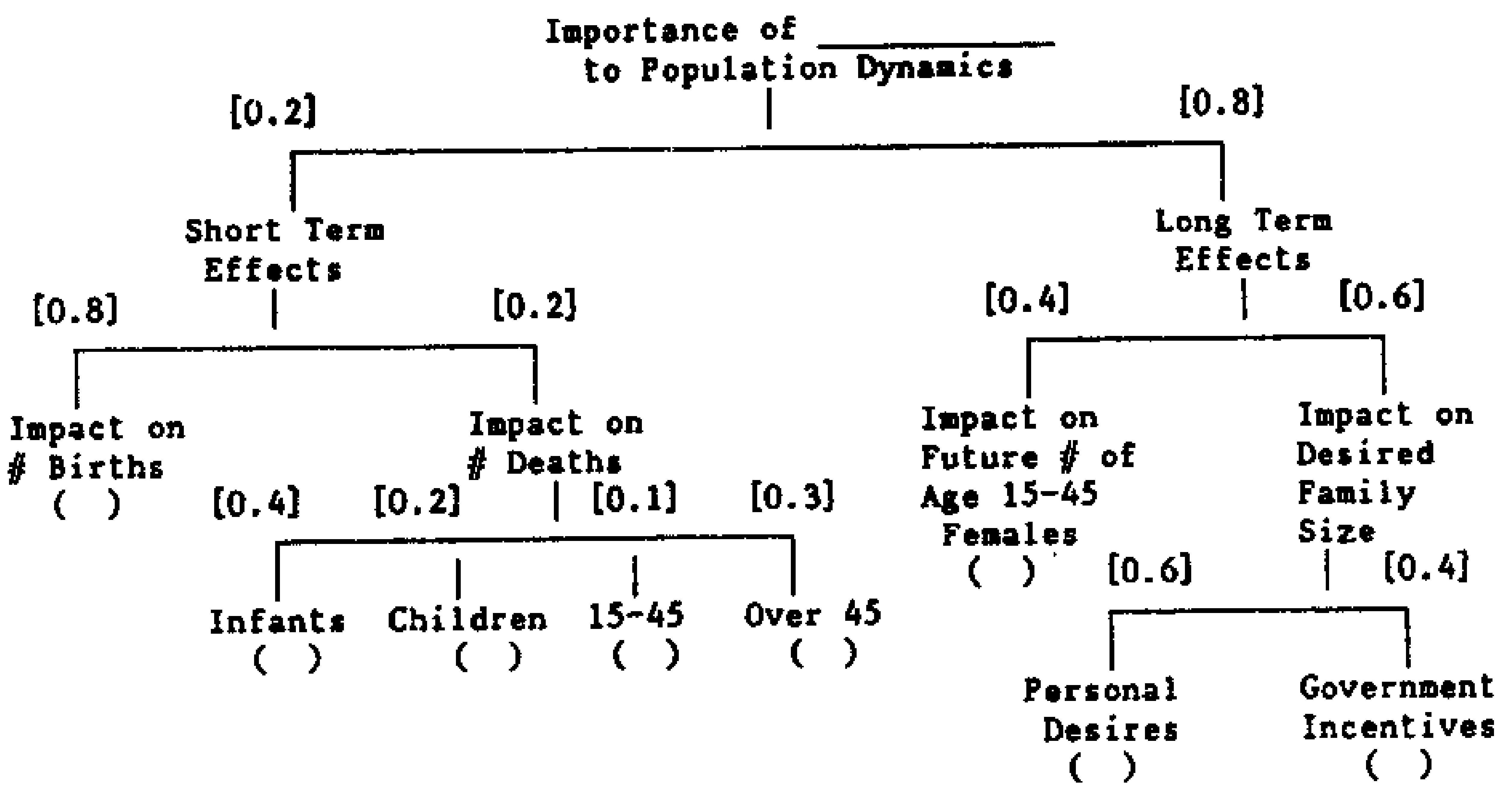
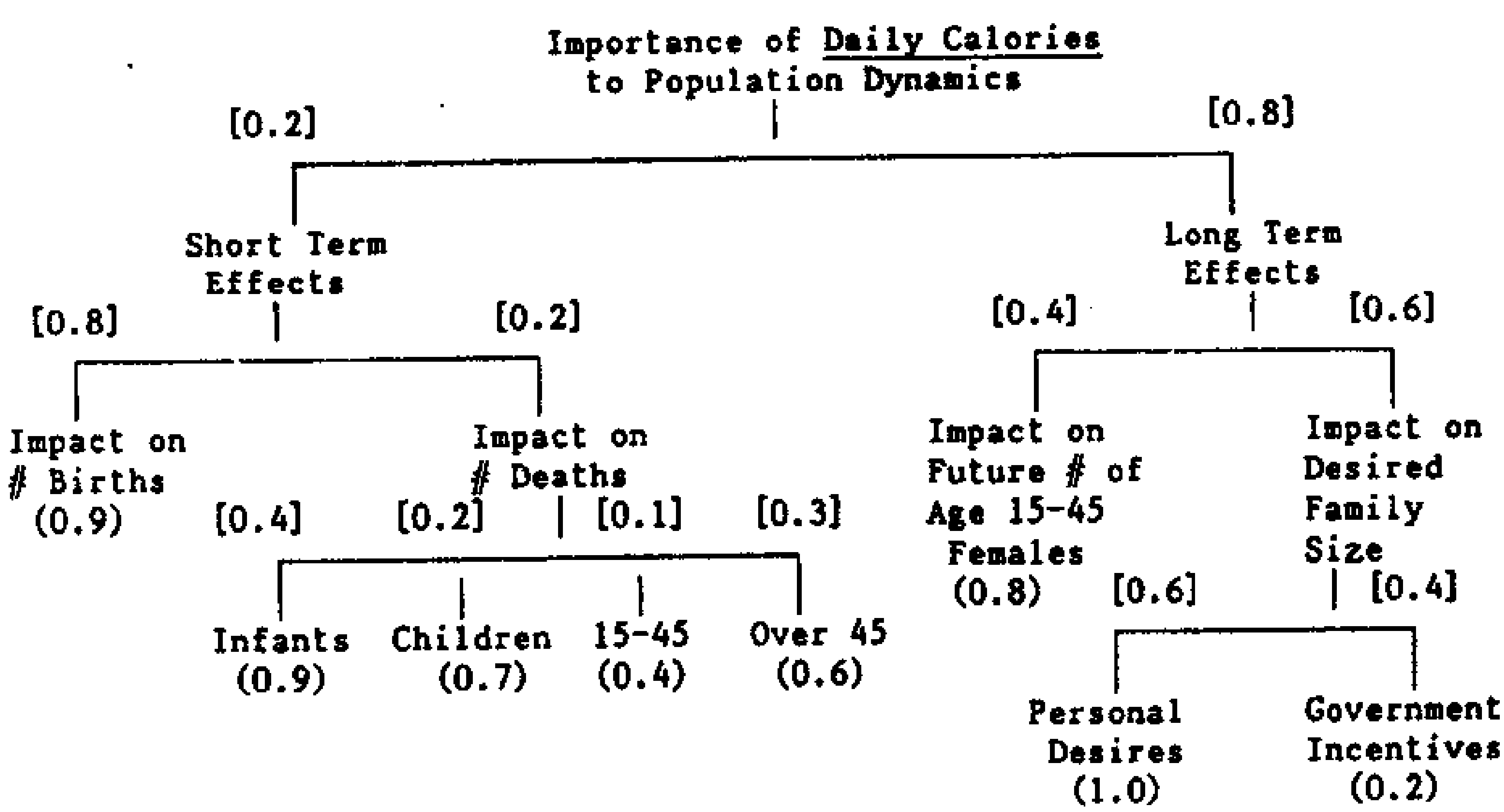


FIGURE 8 -- SAMPLE HIERARCHICAL EVALUATION STRUCTURE (WITH WEIGHTS AND VALUES)



To evaluate the tree we must collapse it from the bottom nodes upward. To do this we multiply each bottom node importance value times its weighting factor, and then sum the resulting products from each branch belonging to a common higher level node. This gives the "value" of the higher level (intermediate) node. For example, the value of the "Impact on # Deaths" node is obtained by:

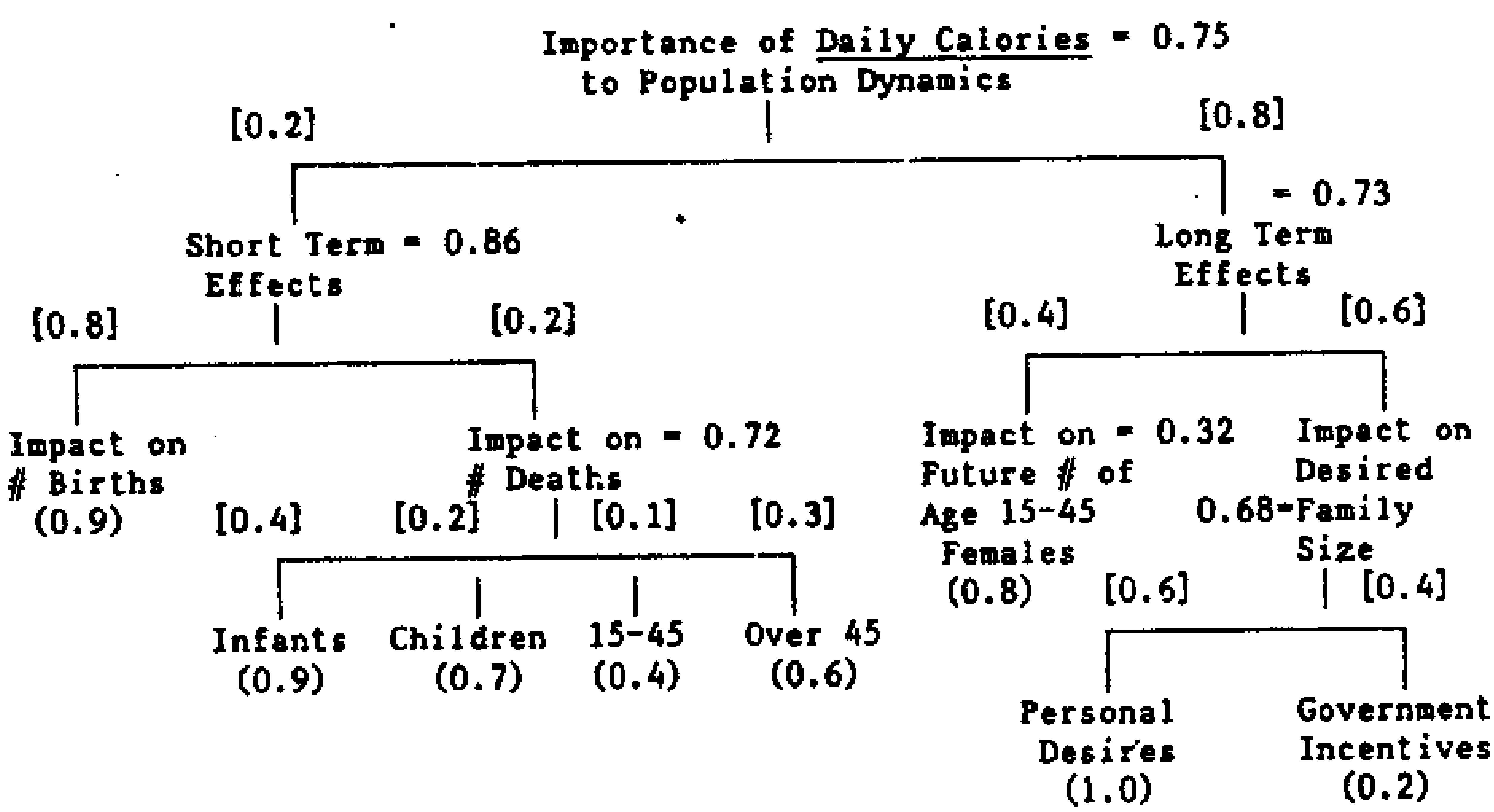
$$(0.9 * 0.4) + (0.7 * 0.2) + (0.4 * 0.1) + (0.6 * 0.3)$$

$$(0.36) + (0.14) + (0.04) + (0.18)$$

$$= \underline{0.72}$$

This value for Impact of # Deaths is then multiplied times its weighting factor and this product is summed with the Impact on # Births value (0.9 * 0.8 = 0.72) to get the 0.86 value indicated below for Short Term Effects. The overall value of Per Capita Daily Calorie Consumption is 0.75. Figure 9 summarizes the results of the calculations.

FIGURE 9 -- COMPLETED SAMPLE HIERARCHICAL EVALUATION STRUCTURE



Additive hierarchical schemes are not the only feasible ones. For instance, if you felt that the importance of the "Impact on Desired Family Size" node would be high (near one) only when both "Personal Desires" and "Government Incentives" were high in value, then a more appropriate calculation scheme might be to multiply together the importance weights given to the bottom nodes. This product would then be used for the value of the higher node. For the values shown in the example above, "Impact on Desired Family Size" would thereby have an importance value of:

$$1.0 * 0.2 = 0.2$$

instead of the 0.68 that it received using the additive logic. Note that the multiplicative scheme does not require weighting factors to be assigned to each branch. The only requirement using the system we've discussed so far, is that the possible values for any given node range from 0.0 to 1.0. Our multiplicative approach meets this standard, since the node being calculated will have a minimum value of zero when either importance score is zero, and will have its maximum value of 1.0 when both importance scores are 1.0.)

FORCE FIELD STRUCTURING

At first glance force field structuring might be considered another variation on the theme of checklists. There is more to the technique than that, however. Instead of listing simply items that impact an intelligence assessment, one tries to list factors that affect some aspect of an analytical problem. Additionally, one indicates graphically the manner in which the factor listed affects the problem aspect under consideration.

For instance, considering again the population dynamics assessment used for illustrative purposes with checklists, one might construct something like the force field representation of factors tending to drive the overall population size toward conditions of under or overpopulation that is shown in Figure 10.

In this example, population size can be seen to be driven (in the eyes of the analyst) toward overpopulation by the factors in the left column, and toward underpopulation by the factors listed in the right column. (If you are doing it by hand the arrows are much more effective at conveying the "force" directionality.) In addition, the display has some inherent use as an analytic tool. In general (we refuse to claim universal truth) things that inhibit movement toward some endpoint are more critical than things tending to drive it in that direction. In the example above, for example, if one feels that the factors births and deaths are of about equal significance, then one would expect population decreases to be achieved more easily by decreasing births than by increasing deaths.

As with checklists, graphical techniques such as weighting the size of the print for each item and/or the thickness of the arrows can be very effective.

FIGURE 10 -- SAMPLE FORCE FIELD REPRESENTATION

UNDERPOPULATION	Driving forces	POPULATION SIZE	Restraining forces	OVERPOPULATION
	Births ------->		<---- Deaths	
	Poverty ------>		<---- Sex Education	
	Increased Life Expectancy ---->		<---- Food Shortages	
	Unemployment-->		<---- Emigration	
	Medical Service Improvements-->		<---- Birth Control	

Disadvantages of Force Field Representations include:

1. The structuring process does not help identify missing factors.

2. Interactions and redundancies among factors are not made explicit.

PROBLEM SOLVING TECHNIQUES

Assessing Information Needs

SEQUENCING TECHNIQUES

Sequencing techniques, sometimes referred to as Time-Line Analysis, involve using time relationships to help establish or examine influence relationships among a set of variables, events, or activities. Basically, and we're giving it only a very basic treatment here, several general classes of use exist:

1. Examining or trying to identify possible cause and effect relationships by looking for fixed time lags between (or virtually simultaneous occurrence of) two events. That is, determining if a common (fixed) time relationship exists between when one event occurs and when another occurs.

   In general this technique cannot prove causal relationships, but it can detect possible ones, and it sometimes can disprove causal hypotheses.

2. Determining the minimum and/or most likely time of occurrence of some event or activity that depends on the prior completion of other, independent events or activities before it can begin.

   This use of sequencing will be covered in some depth as part of the Program Evaluation Review Technique (PERT) section.

Graphical displays probably are used most frequently to look for relationships or to track through an interweaving set of activities, some of which depend on the prior occurrence of some of the others. Tabular structures work well also, but any technique you find useful is certainly acceptable.

PROGRAM EVALUATION REVIEW TECHNIQUE/CRITICAL PATH METHOD (PERT/CPM)

To understand the development and application of PERT (Program Evaluation and Review Technique) as a management device and to apply the technique to a management problem in the field of intelligence.

PERT and CPM (critical path method) are techniques used to aid in the planning and controlling of large, complex projects. Both are often referred to as network-based techniques.

PERT was developed in 1958 to aid in the planning and scheduling of the Navy's polaris missile project which involved over 3000 different contracting organizations.

CPM was developed independently and simultaneously by the Du Pont Company in conjunction with the Univac Division of Remington Rand Corporation.

The two networking techniques are almost identical in concept and in methodology. The primary difference between the two techniques is the manner in which the activity time estimates are arrived at. In PERT, three separate estimates of time are collected for each activity, an optimistic, most-likely, and a pessimistic estimate. CPM requires only a single estimate which would correspond to the most-likely time in the PERT methodology. Because of the similarity between the two techniques, most of the current literature refers to a common name of PERT/CPM. The following construction techniques apply equally to either method.

PERT/CPM networks consist of two major components: activities and events. Activities consume time and resources. Events of the network represent milestones in the project and are represented by a point in time. In constructing the diagram of the model, activities are shown as branches and events are shown as nodes. This procedure was developed because it provides an ease of computation and adaptation to computer modeling packages.

event event
activity

PROBLEM SOLVING TECHNIQUES

Assessing Information Needs

Before the actual diagramming can begin the following steps must be accomplished:

a) a detailed description of all activities that are required to complete the project.

b) determination of the sequence in which the activities must be or will be performed.

c) determination of the time estimated for completion of each activity.

a) The start of a project (which is a point in time rather than an activity) is represented by node 1.

1

b) All activities that are not to be preceded by any other activities start at node 1.

1

c) In general, an activity is to start at node i and end at node j. The activity will be referred to by the name (i,j).

2 (1,2)

1

3 (1,3)

d) The nodes are to be numbered so that if activity (i,j) exists, then $i < j$. (This permits a logical progression through the network).

no yes

4 3 2 4

1 1

2 3

e) All branches leading into a node i represent activities that must be completed before any of the activities leading away from node i can begin.

f) The activity (i,j) must be unique. Again this is to facilitate computations and logical progression through the model.

1 2

Activity A and B would (according rule c) both be referred to as (1,2). To eliminate this confusion a dummy activity is inserted which has no time associated with it:

2

1

3

g) The completion of the project is represented by the highest numbered node.

Definitions

Critical path: The path with the longest time fcr completion. If any activity along this path is delayed, the entire project will be delayed. The slack time for each activity will be zero.

Path Slack Time: The difference between the total time for the critical path and the total time for any other path through the network.

t: The expected time of activity completion.

ES: Earliest activity start time. All predecessor activities must be completed before an activity can start.

LS: Latest activity start time. All following activities must be completed without delaying the entire project.

EF: Earliest activity finish time. $EF = ES + t$

LF: Latest activity finish time. $LF = LS + t$

S: The difference between the latest time an activity can start and the earliest time it can start:

$S = LS - ES$ and $S = LF - EF$

INTERACTION ARRAYS AND DIAGRAMS

The relationships that exist among problem variables can be assessed and displayed in several ways. Interaction arrays and diagrams are among the easiest to use and understand.

The array version of this technique simply uses zeros (0) and ones (1) to indicate the presence (1) or absence (0) of a direct interaction between a pair of problem variables. To develop the array, one lists the problem factors (variables) that are considered relevant along both the left and top margins of a matrix. If there are "n" factors, then a "n" by "n" array will be described by the margin lables. Fortunately, we'll need to fill in fewer than half of the matrix entry positions.

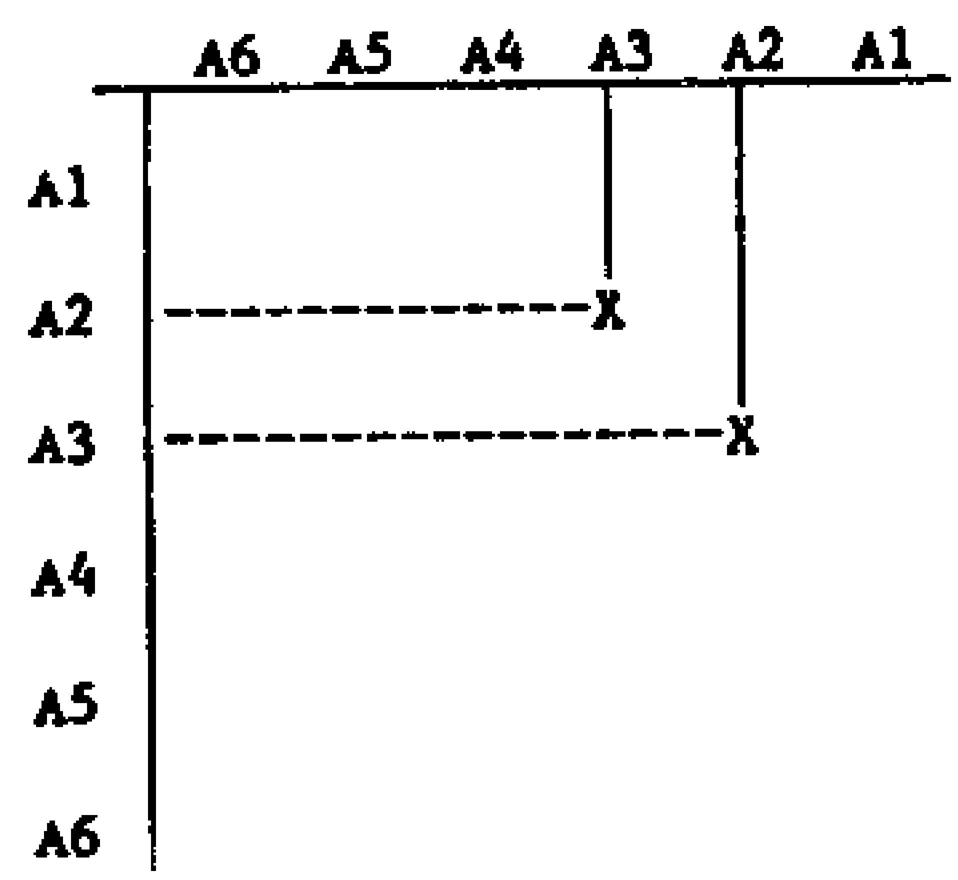
For aesthetic purposes, our final product will look nicer if the top margin lists the factors (from right to left) in the opposite order that they are listed (from top to bottom) on the left margin. For example, assume that our "checklist" of important factors contains six items: A1 through A6. We would structure the margins of our interaction array as shown in Figure 11.

FIGURE 11 -- MATRIX STRUCTURE FOR A SAMPLE INTERACTION ARRAY

	A6	A5	A4	A3	A2	A1
A1						
A2						
A3						
A4						
A5						
A6						

We next enter either a zero or a one in the array locations to indicate whether the two variables, the row and column of which intersect at that point of the array, have a direct association (interaction) or have no direct interaction (indicated by a 1 or a 0 respectively). Except, since the 1's and 0's do not indicate direction of influence, we need fill in only half of the matrix. (Note, for instance, that variables A2 and A3 have in common the two intersections indicated by the "X's" in Figure 12. We need to place a 0 or a 1 in only one of these two intersections.)

FIGURE 12 -- REDUNDANT INTERSECTIONS IN INTERACTION ARRAYS



Also, since each variable does not directly influence itself, we do not need to fill in the array intersections representing the intersection of a variable with itself.

Lets look at an example. Consider the interactions that exist among the variables listed in the left margin of Figure 13. The array shows that the analyst who developed this array felt that, of the factors listed, Oil Prices interact with Agricultural Profit and with the Amount of Fertilizer Used. Agricultural Profit was felt to interact with three additional factors, etc.

FIGURE 13 -- SAMPLE INTERACTION ARRAY

		6	5	4	3	2	1
Oil Prices --------------	1	0	1	0	0	1	
Agricultural Profit -----	2	0	1	1	1		
Crop Yield --------------	3	1	1	1			
Mechanization Level -----	4	0	1				
Amount Fertilizer Used --	5	0					
Inches of Rainfall ------	6						

This array can be depicted graphically. Figure 14 shows an example created simply by drawing a line between those variables having a 1 in their array intersection.

FIGURE 14 -- GRAPHICAL REPRESENTATION OF AN INTERACTION ARRAY

```
       1 ------- 2
      /           \
     /             \
    6               3
     \             /
      \           /
       5 ------- 4
```

To most people, the visual aspects of this diagrammatical display make the interactions easier to identify and understand; but the crossing lines tend to interfere with this improved perception. We can do at least two things to improve on the situation. We can get a little more creative (artistic) with the placement of the factors in the diagram, and we can use the variable names instead of their indentifying numbers. (Figure 15 shows such a revised diagram.)

Note that neither the array or the diagram indicate the direction, magnitude, or other details of the intervariable relationships. They simply identify the existence of relationships. They do not describe them.

CAUSAL ARRAYS AND DIAGRAMS

If we desire to include more detail about variable relationships in our array or diagram, we can use any of several techniques to indicate the direction of interactions that are believed to exist. One basic, plain-vanilla array technique (that we shall call a causal array) uses zeros and ones (as did interaction arrays), but makes use of all of the array intersections

For a causal array, the factors listed in the left margin are considered to cause (affect) the factors listed in the top margin whenever a 1 occupies the corresponding array intersection. Thus, by reading from left to right across a row, we can determine which variables are affected directly by the "row variable" in the left margin. Similarly, by reading down a column we can determine which variables influence the top margin, "column variable" directly.

We shall assume that the factors do not affect themselves, and will thus place zeros along that diagonal. Also, note that since we are now using the entire array, a sense of aesthetics no longer induces us to invert the order in which we list the factors in the top margin relative to their order in the left margin.

FIGURE 15 -- REARRANGED AND LABELED GRAPHS OF AN INTERACTION ARRAY

```
        1
      /   \
     /     \
    2       5
     \     /
      \   /
        4
       / \
      /   \
     6     3
```

```
                 Oil
                Prices
                 /     \
                /       \
   Agricultural          Fertilizer
      Profit                /
          \                /
           \              /
             Mechanization
              /           \
             /             \
            /               \
           /                 \
     Rainfall              Crop Yield
```

An example using the same variables used in the interaction array examples above might look like Figure 16.

In this example, we can determine that Oil Prices are depicted as affecting Agricultural Profit, but that the inverse is not true. There is a 1 in the (1,2) intersection, but a 0 in the (2,1) intersection. Similarly, since both the (2,4) and the (4,2) intersections have 1's, we can determine that "Agricultural Profit" and "Mechanization Level" affect each other.

As with interaction arrays, we can depict this causal array in diagram form also. This time, instead of connecting the factors with lines, we'll use arrows to indicate the direction of the influence. (If two factors each affect the other, we suggest using two separate, unidirectional arrows, rather than one double-headed arrow, since the latter option has less visual impact.)

Using the same aesthetic considerations as above, we would produce the causal diagram shown in Figure 17. (You will need to draw in the arrows for this figure, since the word-processing system used for this text is not very good at making arrows.)

FIGURE 16 -- SAMPLE CAUSAL ARRAY

	1	2	3	4	5	6
Oil Prices ------------- 1	0	1	0	0	1	0
Agricultural Profit ----- 2	0	0	0	1	1	0
Crop Yield ------------- 3	0	1	0	0	0	0
Mechanization Level ----- 4	0	1	1	0	1	0
Amount Fertilizer Used -- 5	0	0	1	0	0	0
Inches of Rainfall ------ 6	0	0	1	0	0	0

For the mathematically inclined, interaction arrays do provide an additional useful function. Using matrix algebra we can determine the directional interconnectedness of the variables. That is, we can determine whether or not a given variable affects another variable indirectly, as well as directly (we indicated only direct influences in the causal array). For small numbers of variables this can be done easily through the use of the causal diagram, but large numbers of variables (particularly when artistic creativity is unable to eliminate crossing lines) sometimes make 3rd, 4th, or more distant interconnectedness less than obvious.

The method for determining directional interconnectedness consists of adding the causal matrix [A] to its identity matrix [I] and then using Boolean multiplication to multiply this summed matrix [A'] times itself until it no longer changes. The final, steady-state matrix will still be (for "n" variables) an n by n matrix consisting of zeros and ones; but the ones in this matrix indicate that a directional influence can be traced from the variable in the left margin to the top margin variable, even if no direct influence exists.

Sample calculations for the interaction matrix shown in Figure 16 are shown in Figure 18.

This final, steady-state matrix allows us to tell at a glance what variables are (or are not) affected by what other variables. Reading from left to right across a row tells us (if a 1 is present) that in the causal diagram one can trace a continuous causal relationship from the left margin variable to the top margin one. If a zero is present, then this causal trace (a causal connectedness) does not exist.

FIGURE 17 -- SAMPLE CAUSAL DIAGRAM

1

2 5

4

6 3

Oil Prices

Agricultural Profit

Fertilizer

Mechanization

Rainfall

Crop Yield

SIGNED CAUSAL DIAGRAMS

Causal diagrams can be used to analyze, as well as structure complicated, interacting systems. To do this, one simply indicates on a causal diagram (and/cr on the equivalent causal array) whether each arrow represents a positive or negative influence. If the affected factor tends to increase as the factor affecting it increases, and also decreases as the influencing factor decreases, then the relationship is positive. A plus sign (+) should be used to lable the arrow. If the two factors change in opposite directions (the affected factor increases when the influencing factor decreases, and vice versa), then the relationship is negative (inverse). A minus sign (-) should be used to lable the arrow in this case.

After assessing each arrow in a causal diagram in this manner, one then can examine the unidirectional loops that are present. If a loop has an even number of minus signs (zero is a even number), then the loop will tend to exhibit uncontrolled (exponential) growth in the absence of other influences. Loops having an odd number of minus signs will exhibit goal

FIGURE 18 -- MATRIX ALGEBRA CALCULATIONS OF CAUSAL CONNECTEDNESS

The summation matrix, [A'], is calculated:

$$\begin{matrix} 0 & 1 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 & 1 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 1 & 1 & 0 & 1 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 \end{matrix} + \begin{matrix} 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 \end{matrix} = \begin{matrix} 1 & 1 & 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 1 & 1 & 0 \\ 0 & 1 & 1 & 0 & 0 & 0 \\ 0 & 1 & 1 & 1 & 1 & 0 \\ 0 & 0 & 1 & 0 & 1 & 0 \\ 0 & 0 & 1 & 0 & 0 & 1 \end{matrix}$$

$$[A'] * [A'] = [A''] = \begin{matrix} 1 & 1 & 1 & 1 & 1 & 0 \\ 0 & 1 & 1 & 1 & 1 & 0 \\ 0 & 1 & 1 & 1 & 1 & 0 \\ 0 & 1 & 1 & 1 & 1 & 0 \\ 0 & 1 & 1 & 0 & 1 & 0 \\ 0 & 1 & 1 & 0 & 0 & 1 \end{matrix}$$

$$[A''] * [A'] = [A'''] = \begin{matrix} 1 & 1 & 1 & 1 & 1 & 0 \\ 0 & 1 & 1 & 1 & 1 & 0 \\ 0 & 1 & 1 & 1 & 1 & 0 \\ 0 & 1 & 1 & 1 & 1 & 0 \\ 0 & 1 & 1 & 1 & 1 & 0 \\ 0 & 1 & 1 & 1 & 1 & 1 \end{matrix} = [A'''']$$

Therefore: [A'''] = [Steady State]

seeking (controlled growth) behaviors.[26]

Frequently one encounters an arrow that seems to need both plus and minus signs, depending on the situation. For instance, does food increase or decrease life expectancy? It depends on whether one currently is under- or overfed. How does one handle this via causal diagramming?

To begin with, this kind of situation indicates a problem exists in the initial causal diagram. That is, by putting plus and minus signs on the arrows, we have identified a weak link in our chain of causal logic. Arrows that seem to have indeterminate signs usually indicate that we have been too simplistic in specifying the nature of the influence the arrow is

---

[26] The goal seeking behavior may take a number of different forms, e.g., damped versus undamped oscillation. The specific form will depend on the exact mathematical relationship of the cause and effect relationship(s).

mean to represent. The indeterminate sign (positive or negative relationship) of the arrow suggests quite strongly that more than one influence exists between the two factors. We often may want to determine them explicitly if we desire to understand more fully the complete system under study. For example, in the food example above, this indeterminate arrow might lead us to the realization that food itself does not influence life expectancy; rather, it is the undernutrition and the overeating that result from the level of food consumption that affect life expectancy.

If a factor belongs to more than one loop, and both positive and negative loops are present, the stronger loop will tend to dominate. But, the diagram will not tell you which loop is stronger (strongest). You must use your analytical judgement for this. Hopefully the diagram will help you recognize that the assessment needs to be made. It also should aid in forming a consistent assessment by clearly defining the factors being assessed.

GRAPHING MATERIAL AND INFORMATION FLOWS

At times it is useful to extend the degree of detail that is expressed in a causal diagram. One method for doing this is flow diagramming.

Flow diagramming involves the use of a set of graphical figures to indicate specific real-world-like activities that affect information or material. For instance, accumulations of material objects (such as the number of tanks in an order of battle) are depicted by a rectangle. Movement of material objects (e.g., dollars from a savings to a checking account) occur along solid arrows that pass through a figurine roughly resembling a valve (which represents the rate of object movement).

Information flows are depicted using broken (dashed) arrows. Circles are used to indicate how information is used (e.g., for decision making or estimating things like degree of discontent).

Other symbols are used also to represent such things as delays in the movement of material objects or information, external factors (e.g., weather) impinging on the system being diagrammed, and sources or sinks of material.

In general, flow diagramming is most commonly used as an intermediate stage in developing mathematical expressions for the relationships believed to describe a system's behavior. However, its use to force ones thinking to a higher degree of detail (even if only on a small part of a problem) is sometimes justified. For instance, this process (since it requires one to think relatively concretely about the processes extant in the real world) may help in the identification of problem related observables against which collection resources can be targeted.

DESCRIPTIVE TECHNIQUES

OVERVIEW OF DESCRIPTIVE STATISTICS

SOME IMPORTANT SYMBOLS

The Summation Sign

The Greek capital sigma is commonly used to indicate summation. It often is referred to as the summation sign. When it appears immediately to the left of an algebraic symbol, it is read as "the sum of." Often, the Sigmas and any variables (e.g., x) have sub- and super-scripts to indicate the numbers of the first and last variable (when they are given in a defined sequence) to be included in the summation. (When the sub- and super-scripts on the summation sign are omitted, the--rather sloppy--indication is that the summation applies to all variable values.)

Multiple Summation Signs

Summation signs often are placed in pairs to indicate the summation of all values in a table (a 2-dimensional array) or in groups of "n" to indicate the sum of all entries in an "n" dimensional array.

MEASURES OF CENTRALITY

The following terms represent various methods for using a single number to indicate what a "typical" or "average" or "middle" value for the population data would be. It might be thought of as the data's center of gravity. Each measure (term) discussed is calculated differently and each can have a different value for the same set of data. Hence, each has a slightly different meaning, and is useful in slightly differing ways and for differing types (metrics) of data.

Mean

Although there are several types of means, we shall use the term only to refer to the "arithmetic mean." This mean is simply the algebraic average of all the observations (data) in a sample population--i.e., the value obtained by summing all of the observations and dividing the total by the number of observations (see Figure 19). The symbol of an overlined "x" (pronounced x-bar) is often used tc represent the mean of a data set (i.e., a sample), and the lower-case Greek letter mu is used for a population mean (i.e., the true parameter value.

FIGURE 19 -- EXAMPLE CALCULATION OF AN ARITHMETIC MEAN

EXAMPLE DATA

	DAY NUMBER									
	1	2	3	4	5	6	7	8	9	10
NUMBER CUPS OF COFFEE	2	4	3	3	6	3	1	3	4	2

$$\text{Mean} = \frac{2+4+3+3+6+3+1+3+4+2}{10} = \frac{31}{10} = 3.1$$

Median

The median refers to the value of the middle ranking observation. That is, if all observations are ranked from the smallest to the largest, the value of the observation with the middle rank is the median. In the case of an even number of observations, the mean of the two middle ranking values often is used as the median value (Figure 20) although, for ordinal data any value lying between the two middle values (inclusive) could be used.

FIGURE 20 -- EXAMPLE DETERMINATION OF A MEDIAN

EXAMPLE DATA

	DAY NUMBER									
	1	2	3	4	5	6	7	8	9	10
NUMBER CUPS OF COFFEE	2	4	3	3	6	3	1	3	4	2

Data listed in rank order:
1,2,2,3,3,3,3,4,4,6

$$\text{Median} = \frac{3 + 3}{2} = 3$$

Mode
The mode is simply the observation value having the highest frequency of occurrence (Figure 21). (Note that the mode may not be unique--e.g., a sample can be bi-modal or tri-modal, or might not have any identifiable mode at all.)

Midrange
The midrange is the value that is midway between the smallest and the largest values in the sample. That is, it is the (arithmetic) mean of the largest and smallest values (Figure 22).

Weighted Mean
The weighted mean is used to calculate the "mean" for grouped data--that is, for data presented in tabular (or lumped together) form. It is obtained by summing the products of the number of observations in each group and the value assigned to each group's observations, and by then dividing this sum of products by the total number of observations in the sample.

If the actual value of every observation within each group is exactly the same, then the weighted mean is equal mathematically to the arith-

FIGURE 21 -- EXAMPLE DETERMINATION OF A SAMPLE'S MODE

EXAMPLE DATA

	DAY NUMBER									
	1	2	3	4	5	6	7	8	9	10
NUMBER CUPS OF COFFEE	2	4	3	3	6	3	1	3	4	2

Mode = 3
(There are more 3's than any other number.)

FIGURE 22 -- EXAMPLE CALCULATION OF A SAMPLE'S MIDRANGE

EXAMPLE DATA

	DAY NUMBER									
	1	2	3	4	5	6	7	8	9	10
NUMBER CUPS OF COFFEE	2	4	3	3	6	3	1	3	4	2

$$\text{Midrange} = \frac{1 + 6}{2} = 3.5$$

metic mean. If, however, each group consists of all observations lying within a certain range of values, then the midrange value of each group is used for the group value and the weighted mean only approximates the true arithmetic mean of the sample.

PROBLEM SOLVING TECHNIQUES

Descriptive Techniques

Modal Class

For grouped data the group having the highest frequency of occurrence is defined to be the modal class.

Median for Grouped Data

The median for grouped data is a special case of the percentile computations to be discussed next, namely, the calculation of the 50th percentile.

PERCENTILES

Data that has been gathered using ordinal or interval scales of measurement can be discussed in terms of the percent of the observations that are greater or smaller in value than a given observation or that have values between two differing value observations. Thus, we can determine that "x" percent of the sample values are smaller than a particular observation of interest, or conversely, we can determine the value below which a certain percentage of the sample observations lie.

For sample data where the exact value of each sample is known, calculating the percentile rank is relatively straight forward. The observations need only to be ranked, and the percent of the observations falling below the observation value of interest gives that observation's percentile rank.

For grouped data the calculations become somewhat more cumbersome. The general mathematical formula for calculating percentile values for grouped data is:

$$\text{Value at the k'th percentile} = b + \frac{(P - F) * w}{f}$$

where:

b = lower boundary for the percentile class

P = kn/100 , n = total number of observations

F = cumulative frequency for all classes lower in value than the percentile class

f = frequency of the percentile class

w = interval width of the percentile class

We also can calculate the percentile rank for a given value "X" from grouped data using the following general formula:

$$\text{Percentile rank for X} = \frac{F + \frac{(X - b)*f}{w}}{n} * 100$$

ANALYSIS ACTIVITIES

Descriptive Techniques

MEASURES OF DISPERSION

Statisical measures of dispersion give an indication of how "spread-out" or how "clumped-together" a set of observations are. Another way of expressing this concept is that these measures provide insight into how much variability (or consistency) there is in the data.

Range

The range is simply the difference between the largest and the smallest elements in a given set of data.

Special ranges can also be given that are based on percentile calculations. For instance, the decile deviation is the difference between the 10th and 90th percentile values. The interquartile range is the difference between the 75th and 25th percentiles. These two percentile ranges give the ranges of the middle 80 % and middle 50 % of the observations, respectively.

Variance

The variance (or its positive square root--the the standard deviation) is the statistical measure most frequently used to indicate the dispersion of sample values of interval or ratio scale data. It is represented by the lower-case Greek sigma squared when referring to populations, and by s^2 when describing a sample. It is calculated by:

1. Finding the difference between the sample mean and each of the sample observations,

2. Squaring each of these differences, and

3. Summing all of these squared values to find their total.

4. Dividing this total by the number of sample observations minus one.

PROBLEM SOLVING TECHNIQUES

Descriptive Techniques

Standard Deviation

The standard deviation is simply the positive square root of the variance. It is represented by the the lower case Greek sigma when referring to the standard deviation of a population, and by "s" when describing a sample.

The Practical Significance of the Standard Deviation

It is now time to introduce a theorem and a rule for using the standard deviation. The theorem was developed by a Russian mathematician, Tchebysheff. Tchebysheff stated that, given a number 'k' greater than or equal to 1 and a set of measurements, at least $(1 - 1/k^2)$ of the measurements will lie within 'k' standard deviations of their mean.

The most important point of this theorem is that the theorem applies to any set of measurements. If we constructed a table based on the theorem it would look like

Values of $(1 - 1/k^2)$

k	$(1 - 1/k^2)$
1	0
2	3/4
3	8/9

so that at least 8/9 of all of the measurements fall within three standard deviations of any mean.

We now state a rule which accurately describes the variability of a bell-shaped (normal) distribution and describes reasonably well the variability of other mount-shaped distributions of data. The frequent occurrence of mound-shaped and bell-shaped distributions of data in nature and hence the applicability of our rule leads us to call it the Empirical Rule.

The Empirical Rule

Given a distribution of measurements that is approximately normal (bell-shaped), about 68% of the measurements will be within one standard deviation (±1s) of the mean, 95% will be within 2 standard deviations, and 99.7% will be within 3 standard deviations.

SCENARIO WRITING

This section has been adapted from a portion of the HANDBOOK OF FORECASTING TECHNIQUES, Part II, Description of 31 Techniques, prepared for the U.S. Army Engineer Institute for Water Resources, Fort Belvoir, VA. The work was done under Contract No. DACW31-75-C-0077, August 1977.

ABSTRACT

Scenario writing is best used to explore possible future conditions, given a set of assumptions. Scenario writing is not well suited either to the making or to the reporting of forecasts. The scenario is essentially the product of the yarn-spinner's art. For serious purposes, however, the assumptions on which the account is based are made explicit. Scenarios have enormous power to communicate the sense or feel of situations which do not and may never exist, a power which is at once an asset and a danger. In a sense, scenario writing is merely a softer version of the well-known modeling, simulation, and gaming methods.

DEFINITION

Originally, scenario referred to the bare outlines of a script for a film or a stage play. As used currently, scenario also refers to an outline of what might happen, given a set of initial assumptions. Thus a scenario is the outline of one conceivable future state of affairs as seen from the general use, perspective writing is an equally pertinent term if the word scenario is considered unclear or unacceptable.

Scenario writing assumes that the major forces or factors which will determine the future state of a given issue are known and can be specified. It also assumes that the author(s) of the scenario are competent to foresee which interactions among forces are most plausible and significant. So long as a scenario is NOT treated as a forecast, these assumptions are reasonable. A scenario writer then becomes nothing more than a plausible yarn-spinner. When the scenario IS treated as a forecast, there is little to support the validity of the assumptions. History suggests that the future, more often than not, will surprise us.

PROBLEM SOLVING TECHNIQUES

Descriptive Techniques

HISTORY

In the broad sense, stories about what will be hereafter are at least as old as ancient religious writings. In modern times, such perspective narratives as George Orwell's 1984 and Aldous Huxley's Brave New World have been powerful influences on conjectures about possible futures. It is important to realize, however, that a scenario may also be a set off numerical projections, based on stated assumptions. EXAMPLE: The U.S. Economy--1995 might well be the title of a scenario consisting of little more than tables of economic, financial, and manufacturing data.

MAIN USES

The most appropriate use of scenario writing is in the situation where it is important to get a feel for what could happen, as opposed to a prediction of what will or may happen. Assume that the Crops of Engineers is suddenly abolished. How would the U.S. be the same and how would it be different ten, twenty, thirty years hence? Often it is difficult to get into a society depicted on the basis of a bare set of assumptions. One needs to play with how the various assumptions relate to and affect each other. Scenario writing is excellent in this application.

LIMITS AND CAUTIONS

As mentioned above, scenario writing can be misleading and counterproductive when it is employed or accepted as an account of the most probable future, rather than as an account of one plausible, conceivable future. Many people who need to know or are interested in knowing what the future will actually be like are too eager to impute an unwarranted confidence in a given scenario. If the yarn-spinner is a skilled craftsman, he may all too easily tap that "willing suspension of disbelief" which is familiar to us all when we read a good book, or view an engrossing film.

OTHER TECHNIQUES

The most appropriate use of scenario writing is to explore complex interrelations among many factors. In that application, perspective writing is most closely related to cross-impact analysis, modeling, simulating, and gaming. In these, too, a major purpose is to learn by experiment, to see how and where and when one change results in other changes elsewhere.

PROCEDURES

Scenario writing is much more an art or a craft than it is a standardized, systematic methodology. For that reason, procedure varies greatly from one application to another. It is fair to observe that scenarios typically are developed by one or at most by a few persons, rather than by large teams as may be the case in, say, survey research methodologies. It is also generally required that the scenario writing begin with a vague, overall image of the situation he wishes to explore. On that basis the writer determines which assumptions he should make, and then selects the nature or the setting of his assumptions. Past that point, statements or components are drafted, reconciled with others, and redrafted, until the final scenario has been completed.

ANALYSIS ACTIVITIES

Descriptive Techniques

PRODUCT OR RESULT

The result is a narrative which may be heavily text, largely numerical-graphic, or a judicious mixture of both. The extent of the scenario statement may range from a few paragraphs to multiple volumes.

A scenario--like a model--must be restricted to deal with only a very few components and interrelations out of many possibilities. The statement therefore typically contains rich detail with respect to the important interrelations among the few components selected for treatment. Other components and details are often ignored or only touched upon in passing.

LEVEL OF CONFIDENCE OF RESULTS

As emphasized throughout in this discussion, scenarios are most appropriately used to explore possibilities, and not to make or express predictions or forecasts. This being the case, the question of confidence in the result is restricted. The only substantive issue concerns whether or not the interrelations described are truly the important ones, and whether they are both logically and plausibly dealt with in the statement.

COMMUNICABILITY OF RESULTS

In this regard, scenarios fare very well compared with other methods. Most people have a great deal experience reading stories, and thus are able to deal easily with complex descriptions of hypothetical future conditions.

CREDIBILITY OF RESULTS

The issue here is a familiar one: How credible is this story? Placed in a literary context, we see at once that credibility of a scenario depends on a number of factors. There is the reader's familiarity with the subject matter. There is the writer's experience and level of craft. There are all the emotional values--accepting and hostile--which attach themselves to the story. Credibility in this case, in short, is largely in the eye of the beholder. It is for this reason among many others that a perspective statement's use is best limited to suggesting and exploring possibilities, rather than to estimate the probabilities.

SPAN OF FORECASTS

As already indicated, scenarios are not well suited to the making or reporting of forecasts. Setting that aside, they are readily adaptable to whatever future time-span interval may be needed. After all, narratives can be and have been written covering micro-seconds at one extreme and many eons at the other.

RESOURCES

The primary resources required here are those of the imaginative craftsman. The one or few people charged with developing a scenario need to immerse themselves in a rich fund of pertinent information. They need to accept for exploratory purposes a wide range of exotic and improbable possibilities. They need to be sensitive to complex and evasive interrelations, and they must have the ability to set these forth clearly and compellingly. Ideally, the scenario will incorporate and be based! logically on a great deal of valid data about the current state of pertinent matters.

CONTEXTUAL MAPPING

This section has been adapted from a portion of the HANDBOOK OF FORECASTING TECHNIQUES, Part II, Description of 31 Techniques, prepared for the U.S. Army Engineer Institute for Water Resources, Fort Belvoir, VA. The work was done under Contract No. DACW31-75-C-0077, August 1977.

ABSTRACT

Contextual mapping is a technique for identifying plausible sequences of development in a given field, and for relating these sequences to potential further developments of particular interest. It has been used largely in technological forecasting applications, although it holds promise in other fields as well. The method is at least as useful for provoking imaginative yet systematic conjecture as it is for predicting actual future developments. Its use requires experienced experts familiar both with the method and with the topic of inquiry.

DEFINITION

Contextual mapping was developed for technological forecasting purposes, and is still used largely in that field, although its use for societal, economic, and other forecasts has been suggested. The principle is simple although its application is often difficult: Any given sequence of assumed technological development will make possible certain further developments. One therefore can specify any development sequence desired and, from it, try to discover what further developments would thereby be enabled and encouraged. Or--to apply the same principle the other way round--one can specify some future technological development of interest and, working backward, try to discover how many different sequences of preceding technological development could lead up to the development of interest. EXAMPLES: (1) One might specify one or more sequence of future developments in solid-state circuitry and, for each, explore what devices or systems might be invented or diffused on the basis of the new technological capabilities assumed. (2) One might specify some generally desirable features for a new high-speed ground transportation system. On the basis of those specifications, one could then try to estimate how many different sequences of RDT&E could enable the development of such a system.

The method assumes that current technological developments will evolve in foreseeable directions. It is assumed that logical, plausible new "clusters" of emerging technological capabilities can be foreseen. It is assumed that the uses to which new technological clusters might be put can be foreseen. It is assumed that future technology-responsive needs can be foreseen, and can be related to emergent sequences of technological developments. It is assumed that the preceding capacities can be seen far enough ahead in sufficient detail to make contextual mapping a profitable investment in support of research and product planning, capital budgeting, etc.

HISTORY

Jantsch informs us that RAND Corporation used the earliest versions of contextual mapping in the late 1940s to forecast the best engine-airframe combinations for U.S. strategic bombers of the 1950s and 1960s. Fairchild Semiconductor relied heavily on contextual mapping in timing its highly successful entry into the integrated circuits field. Many other specific applications, past and current, can be readily found in the literature.

MAIN USES

As mentioned, contextual mapping has been used largely for technological forecasting applications. Donald Pyke, however, reports that TRW has used a modified version of the approach for product planning purposes which considers many non-technological factors, including political, social, cultural, ethical, and ecological considerations. TRW's multilevel map also is tied to a specific time horizon, whereas earlier versions have emphasized date-free sequential maps.

LIMITS AND CAUTIONS

Contextual mapping requires either that its use be limited to topics--e.g., technological development--about which firm estimates can be made or that it be used only to provoke imaginative thought (not predictive forecasts relied upon for decisions). In either case, the technique requires experts well experienced in the topics at hand who also understand the strengths and limits off the method. In many cases, topic focus must be rather narrow in order to use the method within the limits of available time, manpower, and resources.

OTHER TECHNIQUES

Depending on the topic of interest, multiple--regression correlation of trend curves, scenarios of various sorts, and simulation models might yield comparable results.

PRODUCT OR RESULT

The product is one or a series of graphic displays ("maps") depicting sequential events (with or without associated dates). The maps also may indicate what further outcomes a given sequence might be expected to yield. Map data is tied to support documents showing data sources, assumptions made, calculations performed, etc.

LEVEL OF DETAIL OF RESULTS

Where extensive data and information is available about the current developments projected, logical inferences about future developments can plausibly be drawn in great detail. Where such data and information is lacking, details inferred should be regarded with extreme caution but may still be useful to stimulate imaginative conjecture.

LEVEL OF CONFIDENCE OF RESULTS

Confidence levels associated with this method are determined by the level of confidence placed in the intelligence, experience, and judgment of those who apply it. Selection of sequences, identification of potential further outcomes, timing estimates--every major facet of the method hinges on the caliber of judgmental decisions.

COMMUNICABILITY OF RESULTS

Where the topic is sufficiently narrow and only a few well-understood sequences explored, results can be communicated effectively to persons having varied backgrounds and levels of training and experience. Where significant complexities must be assessed, sustained and concentrated attention is required to convey results, even between persons experienced in using the method.

CREDIBILITY OF RESULTS TO CRITICS

If presented as a forecast of some possibilities, results from the method will generally be accepted as credible. If offered as a forecast of the most probable actualities, results are properly open to criticism, since a different set of judgmental decisions at critical points would yield substantially different results.

SPAN OF FORECAST

Typical time spans with this method vary between a few years and a few decades, largely dependent on the normal innovation cycle length in a given sector of activity.

RESOURCES NEEDED

The principal resource required is one or more persons trained and experienced in the technique. Current state-of-the-art data and information is usually needed in some detail, and in typical applications is already available and need only be compiled. A few man-months to a few man-years for a period of many weeks to several months may be required to complete one mapping sequence.

FORECASTING TECHNIQUES

GROUP TECHNIQUES

PANEL CONCENSUS

Using a panel of experts to make analytical judgments is a common method of trying to reach conclusions or to sort through a complex array of interdisciplinary data. The quality of the conclusions reached by such panels vary along a number of parameters: their level of relevant expertise, their motivation to optimize the quality of their output, the clarity with which they understand the actual problem area to be addressed, and the effectiveness of their group dynamics (process).

This last factor, the group process, also impacts (either adversely or favorably) on the other three aspects of group performance, since it can determine the extent to which extant expertise is exploited successfully, the degree of involvement of the panel members, and the correctness with which they identify (or define) the problem statement(s).

Because of the importance of group dynamics to the success of group problem solving efforts, it is worth becoming familiar with those factors contributing to enhanced or degraded group performance. You will be able to affect or control a significant number of the factors discussed below whenever you are a group leader, and can impact them substantially even as a group member.

Assets and Liabilities in Group Problem Solving

Research on group problem solving reveals that the group has both advantages and disadvantages over individual problem solving. The comparison of the merits of group versus individual problem solving depends on the nature of the problem, the goal to be achieved (high quality solution, highly accepted solution, effective communication and understanding of the solution, innovation, a quickly reached solution, or satisfaction), and the skill of the discussion leader. If the potentials for group problem solving can be exploited and if its deficiencies can be avoided group problem solving can attain a high level of proficiency.

The forces operating in groups include some that are assets, some that are liabilities, and some that can be either assets or liabilities, depending upon the skills of the members, especially those of the discussion leader.

PROBLEM SOLVING TECHNIQUES

Forecasting Techniques

Group Assets

Greater Total of Knowledge and Information. There is more information in a group than in any of its members. Thus problems that require the utilization of knowledge should give groups an advantage over individuals.

Greater Number of Approaches to a Problem. Individuals get into ruts in their thinking. Since group members do not have identical approaches, each can contribute by knocking others out of ruts in thinking. Thus group solutions tend to be tailored to fit the interests and personalities of the participants. Group solutions to problems involving fairness, fears, facesaving, etc., tend to vary from one group to another. An outsider cannot process these variables because they are not subject to logical treatment.

Participation in Problem Solving Increases Acceptance. Many problems require solutions that depend upon the support of others to be effective. Insofar as group problem solving permits participation and influence more individuals accept solutions when a group solves the problem than when one person solves it. A low-quality solution that has good acceptance can be more effective than a higher-quality solution that lacks acceptance.

Better Comprehension of the Decision. Decisions made by an individual, which are to be carried out by others, must be communicated from the decision-maker to the decision-executors. Many organizational problems can be traced to inadequate communication of decisions. The chance for communication failures are greatly reduced when the individuals who must work together in executing the decision have participated in making it. They understand not only the solution, but also the other alternatives that were considered and and the reasons they were discarded. A full knowledge of goals, obstacles, alternatives, and factual information is essential to communication. This communication is maximized when the total problem-solving process is shared.

Group Liabilities

Social Pressure. Social pressure is a major force for conformity. The desire to be a good group member and to be accepted tends to silence disagreement regardless of its objective quality and logical and scientific soundness. Problems requiring solutions based upon facts, regardless of feelings and wishes, can suffer in group problem-solving situations.

It has been shown that minority opinions in leaderless groups have little influence on the solution reached, even when these opinions are the correct ones. Reaching agreement in a group often is confused with finding the right answer.

Valence of Solutions. When leaderless groups engage in problem solving, they propose a variety of solutions. Each solution may receive both critical and supportive comments,as well as descriptive and explorative comments from other participants. If the number of negative and positive comments for each solution are algebraically summed, each may be given a valence index. The first solution that receives a positive valence value of 15 tends to be adopted to the satisfaction of all participants about 85 percent of the time, regardless of its quality. Higher quality solutions introduced after the critical value for one of the solutions has been reached have little chance of achieving real consideration. Once some degree of consensus is reached, the jelling process seems to proceed rather rapidly.

The critical valence value of 15 appears not to be greatly altered by the nature of the problem or the exact size of the group. Rather, it seems to designate a turning point between the idea-getting process and the decision-making process (idea evaluation). A solution's valence index is not a measure of the number of persons supporting the solution, since a vocal minority can build up a solution's valence by actively pushing it.

Individual Domination. In most leaderless groups a dominant individual emerges and captures more than his share of influence on the outcome. He can achieve this end through a greater degree of participation (valence), persuasive ability, or stubborn persistence (fatiguing the opposition). None of these factors is related to problem-solving ability.

The mere fact of appointing a leader causes this person to dominate a discussion. Thus, regardless of his problem-solving ability a leader tends to exert a major influence on the outcome of a discussion.

Conflicting Secondary Goal: Winning the Argument. When groups are confronted with a problem, the initial goal is to obtain a solution. However, the appearance of several alternatives causes individuals to have preferences and once these emerge the desire to support a position is created. Converting those with neutral viewpoints now enters the problem-solving process. More and more the goal becomes that of winning the decision rather than finding the best solution. This new goal is unrelated to the quality of the problem's solution and therefore can result in lowering the quality of the decision.

Factors that Serve as Assets or Liabilities

Disagreement. The fact that discussion may lead to disagreement can serve either to create hard feelings among members or lead to a resolution of conflict that produces an innovative solution. A leader can treat disagreement as undesirable and thereby reduce the probability of both hard feelings and innovation, or he can maximize disagreement and risk hard feelings in his attempts to achieve innovation. A skilled leader creates a climate for disagreement which permits innovation without risking hard feelings. The leader's perception of disagreement is one of the critical factors in this skill area. Others involve permissiveness, delaying the reaching of a solution, techniques for processing information and opinions, and techniques for separating idea-getting from idea-evaluation.

Disagreement in discussion may take many forms. Often participants disagree with one another with regard to solutions. Before one can rightly expect agreement on a solution, there should be agreement on the nature of the problem. Even before this, there should be agreement on the goal, as well as on the various obstacles that prevent the goal from being reached. Once distictions are made between goals, obstacles, and solutions (which represent ways of overcoming obstacles), one finds increased opportunities for cooperative problem solving and less conflict.

Often there is also disagreement regarding whether the objective of a solution is to achieve quality or acceptance. When the discussion leader aids in the separation of the several aspects of the problem-solving process and delays the solution-mindedness of the group, both solution quality and acceptance improve. Communications often are inadequate because the discussion is not synchronized and each person is engaged in discussing a different aspect. Organizing discussion to synchronize the exploration of different aspects of the problem and to follow a systematic procedure increases solution quality. The leadership function of influencing discussion procedure is quite distinct from the function of evaluating or contributing ideas. It determines whether a discussion drifts toward conflicting interests or whether mutual interests are located. Cooperative problem solving can only occur after the mutual interests have been established.

Risk Taking. Groups are more willing than individuals to reach decisions involving risks. Taking risks is a factor in acceptance of change, but change may either represent a gain or a loss. The best guard against the latter outcome seems to be primarily a matter of a decision's quality.

Time Requirements. In general, more time is required for a group to reach a decision than for a single individual to reach one. If problems require quick decisions, individual decisions are favored; If acceptance and quality are requirements, group decisions may be preferred.

The practice of hastening a meeting can prevent full discussion, but failure to move a discussion forward can lead to boredom and fatigue-type solutions, in which members agree merely to get out of the meeting. The effective utilization of discussion time (a delicate balance between permissiveness and control on the part of the leader) is needed to make the time factor an asset rather than a liability. Unskilled leaders tend to be too concerned with reaching a solution and therefore terminate a discussion before the group potential is achieved.

Who Changes. In reaching consensus or agreement, some members of a group must change. If persons with the most constructive views are induced to change the end-product suffers. The leader can upgrade the quality of a decision by protecting the person with a minority view and increasing his opportunity to influence the majority position. This protection is a constructive factor because a minority viewpoint influences only when facts favor it. Leaders who see some of their participants as troublemakers obtain fewer innovative solutions and gain less acceptance of decisions made than leaders who see disagreeing members as persons with ideas.

ANALYSIS ACTIVITIES

Forecasting Techniques

SOME OBSERVATIONS ON USING INTERDISCIPLINARY GROUP PROBLEM SOLVING

KEY JUDGEMENTS[27]

Awareness of the inability of a single discipline to provide meaningful answers to issues that cross disciplinary lines has led to an increasing stress on interdisciplinary team research. Team research is not always successful, however, because it is difficult to effectively integrate insights, concepts, and research methods from several different disciplines. As a result, numerous evaluation studies of interdisciplinary team research projects have been conducted to determine why some projects succeed and others fail. Judgements drawn from those studies that are relevant to the organization and management of interdisciplinary research in the Agency are summarized here.

<u>Appropriateness</u>

--Interdisciplinary team analysis (ITA) is <u>not</u> appropriate for most intelligence problems.

--ITA should be reserved for very complex problems that involve three or more disciplines and that require a high level of interaction between team members over an extended period of time.

<u>Personnel Considerations</u>

--Not all analysts make good team players. Selection of the team leader and team members is a critical stage in the process. Likely candidates include analysts familiar with more than one discipline and/or those who view the ITA experience as an opportunity for expansion. Team members should be "discovery minded and problem oriented rather than status minded and reward oriented."

--ITA is most successful when the team leader is both substantively expert and a good manager; least successful when he/she is a poor manager.

--Team size in successful ITA projects has varied considerably. Most reviewers conclude that in large teams (over 10 members) originality declines and interpersonal problems increase.

--ITA is more likely to succeed if all the key participants are available at the beginning of the project.

--The longer the team stays together, the more successful it becomes.

----

[27] This section on Interdisciplinary analysis is taken from the Key Judgements section of a paper, "Interdisciplinary Team Analysis: Implications for Managers," written by Fred Grupp for the Office of Research

PROBLEM SOLVING TECHNIQUES

Forecasting Techniques

Critical Phases

--Problems are most likely to develop during the problem formulation, research design, and problem interpretation (task assignment) phases of the project. Papering over of differences in these early stages of the analysis inevitably leads to long-term failure.

--A clear statement of the problem assists the research team, but does not substantially speed-up the process. Theoretical compatibility is essential for successful ITA and this is rarely achieved quickly.

Support Requirements

--The strong, consistent and continuing support of top managers is required to overcome the insecurities of team members who see that their work is outside the mainstream of Agency analysis.

--ITA has been most successful when team members were located together, had their own private workspace, and had access to a group meeting room.

--More support infrastructure is required for ITA projects than for single discipline studies because team members are obliged to deal with unfamiliar data sources and methods. The per-capita investment in research assistants and computer support personnel is higher for ITA projects than for individual analysis.

Costs and Benefits

--ITA is time-consuming and expensive. It is likely to identify gaps in collection and lead to the development of new requirements. Methodologies, data bases, and work habits are all likely to require revision as work progresses. Planning and scheduling are made more difficult by these uncertainties. Per-capita efficiency is less for team research.

--ITA can make progress on very complex problems that could not be successfully completed by an individual.

--Team research produces more accurate solutions, but in a less-timely fashion.

ANALYSIS ACTIVITIES

Forecasting Techniques

FUNCTIONS IN TASK ORIENTED GROUPS

Adapted from Penne, K., & Sheats, P. Functional roles of group members. The Journal of Social Issues, Spring, 1948, 42-27; and from Bion, W.R. Experiences in Groups. 1959.

Behavior in a group can be looked at from the point of view of purpose or function. When an individual says something, is he/she primarily trying to get the group task accomplished, or trying to improve or patch up relations among members, or pursuing some personal goal without much regard for the group's problems?

As the group grows in maturity and the needs of individual members (hopefully) become integrated with group goals, there is the expectation that there will be less self-oriented behavior and more task and maintenance behavior.

A. Types of behavior relevant to the group's work on its task:

1. Initiating: Proposing tasks or goals; defining problems; suggesting procedures or ideas for solving a problem.

2. Information-seeking: Requesting facts; seeking relevant information about group concerns; asking for expressions of feeling; requesting statements of estimates; seeking ideas.

3. Information-giving: Offering facts or opinions or perceptions of group concern; stating beliefs or values' making suggestions; offering ideas to solve problems.

4. Clarifying or elaborating: Interpreting ideas and suggestions; clearing up confusion; defining terms; stating the issue.

5. Summarizing: Pulling together related ideas and suggestions; offering conclusions for group acceptance.

6. Consensus-testing; Asking to see if the group is near a decision; sending up a trial balloon to test conclusions.

B. Types of behavior relevant to the group remaining in good working order; having a climate conductive to work accomplishment; and maintaining good relations among members to permit maximum use of member resources, e.e., group maintenance.

1. Harmonizing: Attempting to reconcile disagreements; reducing tension; getting people to explore differences.

PROBLEM SOLVING TECHNIQUES

Forecasting Techniques

2. Gate-keeping: Helping to keep internal group communication channels open; facilitating the participation of others; suggesting procedures that permit greater shared participation.

3. Encouraging: Being warm, friendly, and responsive to others; indicating by facial expression or remark the acceptance of others' contributions.

4. Compromising: When one's own idea or status is involved in a conflict, offering a compromise which yields status, admitting error, modifying position in the interest of group cohesion or growth.

5. Standard-setting and testing: Testing whether group is satisfied with its procedures or its standards for accomplishment; pointing out group norms so they can be reevaluated.

Even aside from such difficulties as the quality of ideas, basic incompatibilities, or incompetence of the group or individual members, not all behavior in problem-solving, task-oriented groups is conductive to group effectiveness. There is an underworld of emotional current and play in the life of most groups, especially as they are beginning to mature or as they face crisis seemingly too severe for their present maturity. These behaviors need to be recognized for what they are: they cannot be ignored or wished away or subjected to judgements or incantations. After recognition, their causes need to be understood.

There are several characteristic causes of self-oriented behavior of the sort with which we are concerned:

1. The problem of identity: Who am I in this group? Where do I fit in? What kind of behavior is acceptable?

2. The problem of goals and needs: What do I want from this group? Are my goals compatible with those of the group? What do I have to offer to the group?

3. The problem of power, control, and influence: Who will control what we do? Is his/her influence on me acceptable, to be trusted? How much power and influence do I have? \How much do I want?

4. The problem of intimacy: How close will we get to each other? How personal is it akay to be in this situation? How much can we trust each other? How can we achieve a greater level of trust?

C. What kinds of behavior are likely to be produced in response to these kinds of problems?

1. Dependency/counterdependency: Leaning on or resisting (to an unrealistic extent) anyone in the group who represents authority, especially the boss (or trainer).

2. Fighting and controlling: Asserting personal dominance, attempting to get one's own way regardless of others.

3. Withdrawing: Trying to remove the sources of uncomfortable feelings by psychologically leaving the group-- or by physical departure, in rare cases.

4. Pairing: Seeking out one or two supporters and forming a kind of emotional subgroup in which members try to protect and support each other.

These are not the only things which can be observed in a group. The main point, however, is that improving our skills in observing what is going on in the group will provide us with important data for understanding groups and increasing our effective leadership with them.

GUIDELINES FOR REACHING GROUP CONCENSUS

The method of "group consensus" in reaching decisions can be beneficial in situations where there is no clear expert and where relevant information is held by different group members.

Group consensus can be a good decision-making technique for using all the available resources in the group and for resolving conflicts in an innovative manner. Complete unanimity is not desired, nor is it often achieved. However, each group member should come to accept the group decision in terms of logic and practicality.

The following guidelines can be used to attain consensus.

1. Avoid arguing for your own decision or choices. Present your position as logically, lucidly, and succintly as you can. However, pay attention to the reactions of the group, think about them, and use them when restating or pressing your opinion.

2. When a stalemate is reached in discussion, don't get into a win-lose confrontation where someone must win at someone else's expense. Instead, find the next most acceptable strategy or alternative for the contenders which is also acceptable to the group.

3. Don't change your mind just to avoid conflict or disagreement for the sake of harmony. When agreement seems to come to quickly, be skeptical, explore the reasons why, and make sure that everyone accepts the solution for similar or complementary reasons. Yield only to positions that seem to you to be objective and logically sound.

4. Do not use techniques such as majority vote, coin tosses, averaging, or bargaining to reduce conflict. If hold-out members finally agree, neither punish them for holding our or reward them

for giving in (e.g., by letting them have their way on some other point).

5. Differences in approach and opinion are natural and to be expected. Seek them out and try to have everyone participate in the decision making. Disagreements can help promote better decisions; with a range of opinions and a variety of information, there is a greater probability of hitting upon a more innovative solution.

ANALYSIS ACTIVITIES

Forecasting Techniques

DELPHI

DELPHI is a systematic method that uses anonymous responses to a questionnaire to help obtain an objective concensus from initially divergent expert opinion. One objective of the DELPHI method is the encouragement, rather than the suppression, of conflicting or divergent opinions. It arrives at a concensus by pooling the two separate items involved in any estimate: 1) Expert information or knowledge, and 2) Good judgment, analysis, and reasoning. Although a DELPHI participant may not initially be well informed on a given question, he still can contribute judgment, analysis, and reasoning of the information and arguments provided by other respondents.

DELPHI is useful in obtaining answers to questions of judgment, for which there is no known or factual answer. This difficult kind of question is common in the Intelligence Community. Estimates and assessments are in the final case questions of judgment. They are based on incomplete facts and inexact knowledge, to which analysis and judgment must be applied. The use of DELPHI results in definite conclusions being reached by groups of experts dealing with incomplete information.

Like a simulation model, DELPHI proceeds in stages or rounds. Each round makes use of the results (information) from preceding rounds to reevaluate and refine judgment. The key element to DELPHI is that the participants contribute their ideas and arguments anonymously. Thus, many of the potential problems of group problem solving techniques are avoided (e.g., the Boss cannot dominate, divergent ideas are encouraged since everyone must respond before the responses of others are known). But creative solutions are unlikely, since the formal polling procedures used prevent the group from easily restating the problem (questions) or diverting to new, but relevant topics.

With this background in mind, let's examine in more detail the DELPHI process itself. We'll make this examination by discussing the DELPHI process in the same order that it progresses; that is, by discussing each round, or iteration, sequentially.

<u>Round One</u>.

<u>Coordinator Preparation</u>. Prior to the beginning of the first round the DELPHI coordinator must prepare the questionnaire that will be used to poll the participants. Participant codes, or some other mechanism for maintaining anonymity, must be developed.

The questions on the questionnaire should be constructed in a manner that requires quantifiable answers (at least ordinal scale). The questions should not permit a bianary choice (e.g., yes-no, a or b,etc.).

<u>Participant Activities</u>. The first round involves each participant receiving and responding to a questionnaire. Best-guess answers are provided to each of the questions and these responses are returned (anonymously) to the DELPHI coordinator.

PROBLEM SOLVING TECHNIQUES

Round Two.

Coordinator Activities. After receiving the participant responses to the first round questionnaire, the coordinator must compile the results and prepare a summary to give to each of the participants for round two. The summary should include either an exact listing of the responses received during round one, or (if the group is large) statistical data (e.g., mean or median, range, standard deviation) for each of the questions.

Each participant is given this summary, and, in addition, is also tasked to provide rationale for each response that he (she) gave during round one that was in either the upper or lower end (e.g., quartiles or deciles--depending on group size) of the composite group answers.

Participant Activities. Having received the summarized results from the round one responses, each participant again answers each question, and, for those questions for which he (she) was in the upper or lower end of the group based on round one responses, provides rationale for why the he (she) answered as he (she) did.

Round two gives each participant the opportunity to argue for specific positions that differ from the group norm in order to persuade fellow participants. This provides an opportunity for participants to influence the group decision [analytic] process.)

Round Three

Coordinator Activities. Summaries of the participant responses are again prepared, as well as a list of the rationale statements received. Both the round two summaries and the list of rationale statements are provided to the participants to help them re-assess their answers for round three. Participants whose round two answers were in the upper or lower extremes of group responses should again be tasked to provide rationale statements. In addition, each participant should be tasked to indicate the rationale statement for each question that is evaluated by that participant to be the least convincing.

Participant Activities. Participants use the summarized data from round two and the list of rational statements to re-assess their answers to the questionnaire. They again provide rationale statements for those answers that lay in the "lunatic fringe" of the group responses. Assessment of the rationale statement that is considered least convincing should be made prior to re-answering each question.

(Round three gives participants the opportunity to influence the group process by judging the arguments of the other participants.)

Rounds Four, Five, etc.

Coordinator Activities. Each round now becomes a repeat of the corresponding round three activity, until a satisfactory group concensus is reached, or it becomes obvious that the group has polarized or in some other manner gotten into a box from which a group concensus cannot be reached. When either of these extremes has been reached the DELPHI session is terminated.

Participant Activities. Participants continue, as in round three, to indicate the least convincing rationale statements and to re-assess their answers based on the new data summaries received each round and their evaluation of the worth of the rationale statements provided by the "lunatic fringe."

PROBLEM SOLVING TECHNIQUES

SAMPLING TECHNIQUES

FUNDAMENTALS

Purpose of Sampling

When one needs to know something about the parameters of a large population (e.g., the mean age of the Iranian population) it frequently is difficult or impossible to make an exact measurement on each individual member of the population. We often are forced, therefore, to make estimates based on a small subset of the whole population. (Sounds rather similar to inferential statistics, doesn't it?) In fact, we'll make extensive use of basic statistical concepts.

Sampling theory makes the estimative process more efficient. It does this by suggesting strategies for selecting the subset to be used in making an estimate, by defining alternative methods for making estimates from the subset data, by helping to reduce the cost of taking samples, and by providing techniques for quantifying (in a predictive or guesstimative sense) the size of the estimative error associated with a given sampling or estimation algorithm.

Advantages of Sampling over Complete Census

In general, sampling techniques enable surveys to be conducted with reduced costs, greater speed, broader scope, and acceptable (sometimes even improved) accuracy relative to complete population surveys. These advantages accrue from the ability of sampling theory to specify, in advance, the size that the sample must be to achieve the accuracy needed, to suggest the most cost effective survey method and sample composition, to put specialized equipment and highly trained personnel to effective use, and to reduce to a minimum the size of the required survey staff.

Improved accuracy relative to a complete count? This may seem contradictory. However, when confronted with extremely large populations and no good automated method to count or measure, boredom, fatigue, and shear numbers can sharply degrade both the accuracy and reliability of whole population counts. Try, for instance, repeatedly counting the number of grains of sand held by a one-cup measure. If you get the same answer twice you're doing very, very well.

ANALYSIS ACTIVITIES

Forecasting Techniques

Steps to Making a Survey

The following provide a guide to the steps involved in making a survey. Some flexibility exists in the exact sequence in which the steps can be accomplished; but they provide a good reference framework and checklist against which one can evaluate the adequacy of a specific survey strategy.

A. State the survey objectives (including the acceptable chance of error--the alpha level--and the accuracy needed.

B. Define the exact population to be sampled.

C. Determine specifically the amount and type of data that is to be collected.

D. Decide on the methods that will be used to make data measurements (e.g., question order, exam or opinion poll).

E. Choose the sampling unit (the things that will be considered to be individual members of the population). (For instance, are you sampling individual soldiers or artillery batteries?)

F. Select the sample (e.g., its size, the population strata it will include, etc.).

G. Organize the field work. (Write your specific collection requirements, manage the collection assets under your control, provide concurrent critique of the early results, etc.)

H. Analyze the data. (Eliminate or correct obvious errors. Organize the data. Make the necessary calculations. Etc.)

I. Make the population estimates (both best-guess projections and error bounds at the specified confidence levels).

J. Prepare for future sampling studies. (Note what went wrong, what could have gone more smoothly, what was done well, etc.)

Sampling Requirements

To use Sampling Theory, one technically must have a <u>probability sample</u>.

Simply defined, a probability sample is a uniform random sample. That is, each of its population members has a known (usually equal) chance of being selected for data measurement. When population subsets exist, they too have known, usually equal chances of selection.

Although this definition may seem restrictive in real world terms, there actually is quite a bit of flexibility in how one defines the population members. For instance, if one wants to survey an adversary's ICBM installations, it might be advantageous to take a proportionately larger sample of those silos bearing missiles with a larger throw-weight capacity or a longer range. In these cases one might define the population members being sampled to be the "Kilograms of throw-weight" or the "Kilometers of range" of each missile, respectively, rather the the missiles themselves.

Similarly, one can break the initial population down into nonoverlapping subsets, and sample each of these "probability samples" differently. This technique of forming population subgroupings is used frequently to conduct stratified population surveys. (Stratified sampling is discussed in more detail below.)

Probability samples have some convenient properties (and Sampling Theory exploits them to the fullest). One of the most important is that the standard error of the sampling method can be calculated. Another important property is that the estimative procedures that can be used are clearly defined and produce a unique estimate (which, because of the first point, can be error bounded at any chosen confidence [1-alpha] level).

SOME SAMPLING STRATEGIES

Simple Random Sampling

Simple Random Sampling is the basic, random draw type of sampling to which all basic statistic classes are exposed. It simply samples the population, as a whole, equally. All population members have an equal chance of being selected.

Stratified Sampling

Stratified samples divide population units into subgroups based on some survey-relevant characteristic(s). Several basic types of stratified sampling procedures exist:

a) Equal-size: an equal sized sample is taken from each subgroup, regardless of the relative sizes of the subgroups.

b) Proportional: the size of the sample taken from each subgroup is proportional to the relative size of the subgroup.

c) Optimum Allocation: for a given total sample size (or cost), the size of the sample taken from each subgroup will be allocated so that the smallest overall error in the estimate (the greatest precision) will be obtained for the specified cost or total sample size.[14]

Systematic Sampling

Systematic sampling uses a patterned, serial, or chain method to select sample units. For instance, every 4th person is polled.

This type of sampling can lead to large biases in the sample data if the characteristic being measured is distributed nonrandomly in order, over time, location, etc. It is extremely dangerous to use this form of sampling if the final sample size is small and access to population members is not random.

Interpenetrating Replicate Subsample Sampling

Another, perhaps clearer, term for this type of sampling strategy is Independent Sample Sampling. What this form of sampling amounts to is taking independently several sets of samples from the same population. These independent (redundant) samples can then be used to check the consistency between (among) the samples (which have been made by different investigators or instruments).

This sampling strategy is particularly useful in estimating the precision of complex sampling procedures.

Two-Phased Sampling

Two-phased sampling (often called Double Sampling) refers to methods that take two separate and independent samples sequentially, such that the details of the second sampling activity are based on the results of the first.

The first sample may be small or large in size, but is broad in scope. It is used to determine, in as fast and inexpensive a manner as is possible, the most relevant characteristics of the population to study in more detail. The second sample follows up on these determinations in sufficient detail to reach the degree of precision required. It usually is (but does not have to be) smaller in size than the first sample, but it will be much more limited in scope, with more intense examination of the characteristics being studied.

This sampling strategy is particularly useful when one does not know for sure, or has no idea of what the question(s) should be. For instance, country observers report that they sense a change in popular support for the current government, but are unsure of what is causing it. In this case, one might what to do a quick, cursory survey that examines all possible causes, and then follow-up with a larger, more detailed survey to

---

[24] This strategy uses a mathematical algorithm that weights the subgroup sample sizes proportional to their strata size and variance, and inversely proportional to their per unit sampling cost.

double-check (confirm or refute) the potentially interesting results indicated by the first survey.

Sequential Sampling

This technique frequently is used when the sample results will drive a decision choice. The basic design is takes a series of small, independent samples sequentially, and examines the results of each sampling "round" prior to starting the next sampling process. Another sample is taken only if sufficient precision to make the decision has not yet been obtained from all of the preceding samples. If the very first sample happened to provide sufficiently conclusive results for the decision to be made, then the entire sequential sampling process would terminate after only one sample had been taken.

This form of sampling is common in acceptance testing and quality control situations.

Sampling Proportional to Size

This techniques gives a higher probability of selection to population members having more importance to the results of the survey. For instance, if one was sampling a country to determine how the next election would go, it might be wise to select more frequently people from cultural strata that votes in higher percentages than people from other strata. Note that the increase in the probability of sample selection needs to be specified clearly to maintain the sample as a probability sample.

In essence what this sampling strategy amounts to is a simple random sample of a population having a redefined sample unit (such as the missile example used earlier in the Sampling Requirements section).

Cut-Off Sampling

This method eliminates from possible selection the smallest valued population units. This can be used and still give the overall required precision in population estimates when the total value of the small-valued population members is itself small. (Note that this is not strictly a probability sampling technique, but it can be much less expensive and less time consuming.)

Combined Sampling Methods

Most sampling procedures used in the real world represent some mix of the techniques available. This is completely legitimate to do as long as a probability sample is used.

ESTIMATION TECHNIQUES FOR POPULATION VALUES

Blow-up Estimates

This method simply multiplies the obtained sample mean times the total population size:

$$\bar{X} * N = \text{Total Population Value}$$

For example, if the mean number of trucks produced per assembly line was found to be 48 trucks per day based on a sample of 15 assembly lines, then the total number of trucks produced per day from all assembly lines (let's assume there are 100 of them) would be 48 * 100 = 4800 trucks per day.

Ratio Estimate

This method requires data from a previous, relevant time period (the reference time period) for both a comparable sample and the actual population value for that reference time, in addition to data from a current sample. By then assuming that the percent error that was determined to have occurred during the first sampling period will occur again, one can estimate the current population value by multiplying the true population value for the reference time period by the ratio of the current sample value to the reference sample value. (Figure 23 shows a graphical representation of this mathematical logic.) The equation form looks like:

$$\text{Population Value} = \frac{\text{Current Sample Value}}{\text{Reference Sample Value}} * \text{True Reference Population Value}$$

Difference Estimate

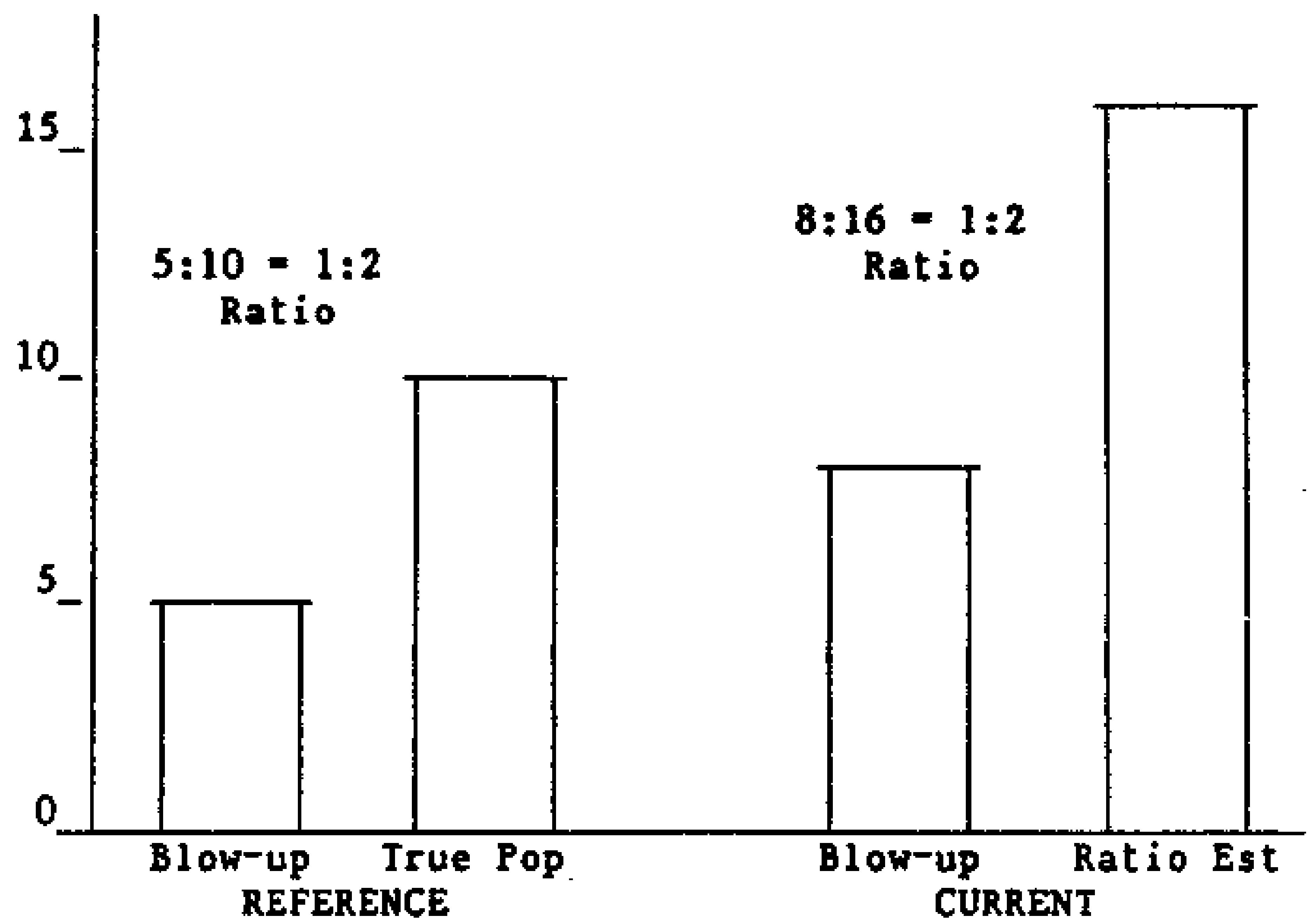
This method assumes a constant absolute error for both the estimate made for some reference time period and for the current time period. Like the ratio estimate, it uses (requires) data from a reference time period in the form of a sample value for that period and the true population value for that same period. It also uses a current sample value.

Calculations are made by making a standard blow-up estimate for the current time period and then adding (algebraicly) to it the error between the blow-up estimate and the true population value for the reference time period. This is illustrated graphically in Figure 24. The equation is:

$$\underset{\text{(now)}}{\text{Difference Estimate}} = \underset{\text{(now)}}{\text{Blow-up Estimate}} + (\underset{\text{(reference)}}{\text{True Pop Value}} - \underset{\text{(reference)}}{\text{Blow-up Estimate}})$$

FIGURE 23 -- GRAPHICAL REPRESENTATION OF RATIO ESTIMATE

Graphical Representation of Ratio Estimate



The RATIO for both time frames remains the same.

Linear-Regression Estimate

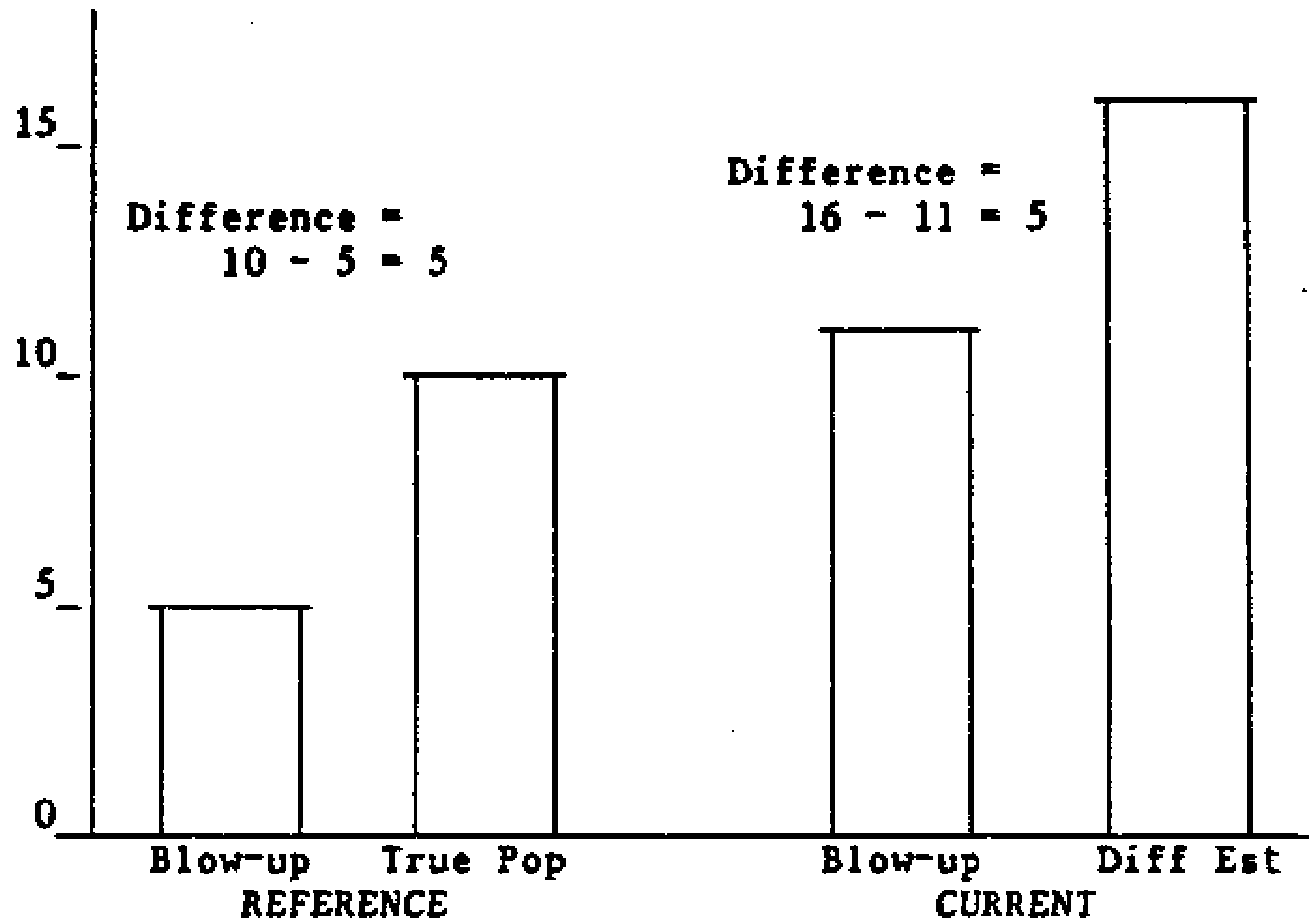
Linear-regression estimates use the regression coefficient calculated between the current and reference time period sample data to "adjust" the correction factor used in the Difference Estimate procedure just discussed. The mathematical procedure looks similar to the difference estimate equation (see Figure 25), except that the regression constant is multiplied times the difference correction factor. This technique is particularly appropriate when the linear regression line determined for the data does not intercept the origin. Using "r" to represent the regression coefficient, the Linear-Regression equation looks like:

Linear-Regression Estimate =
(now)

Blow-up Estimate + [r * (True Pop Value - Blow-up Estimate)]
(now) (reference) (reference)

FIGURE 24 -- GRAPHICAL REPRESENTATION OF DIFFERENCE ESTIMATE

Graphical Representation of Difference Estimate



The DIFFERENCE for both time frames remains the same.

Comparison of Estimate Method Accuracies

In general, the percent error of the estimative methods discussed above is:

Linear-Regression > Difference > Ratio > Blow-up

AFFECTS OF A CONSTANT SAMPLING BIAS

If a sampling procedure introduces a constant bias (b) to be superimposed on the value reported for each datum, the result will produce an "unbiased" estimate around an estimated population mean (u') that has been shifted from the true value of the population mean (u) by the magnitude of the sampling bias.

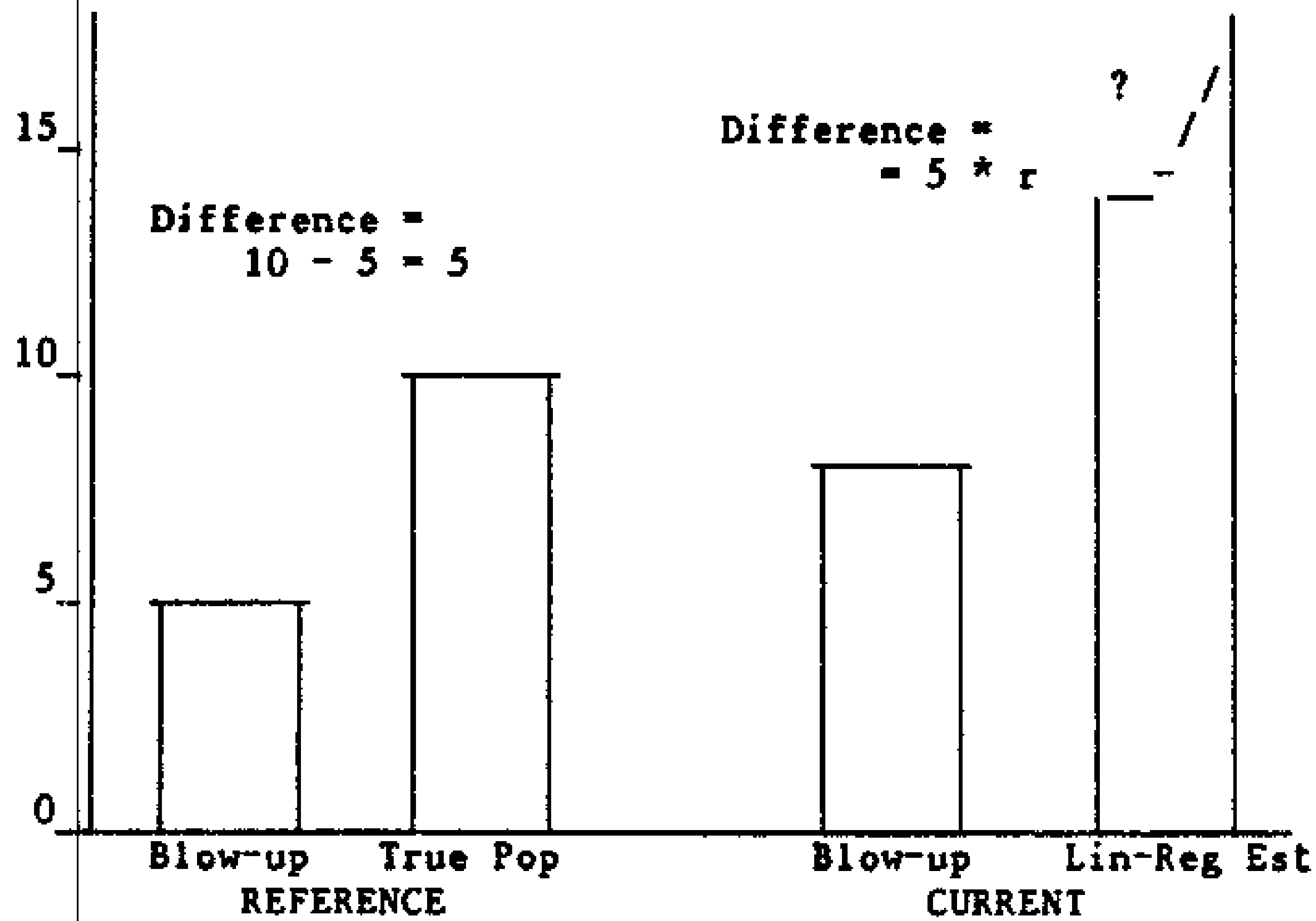
$$u' = u + b$$

PROBLEM SOLVING TECHNIQUES

Forecasting Techniques

FIGURE 25 -- GRAPHICAL REPRESENTATION OF LINEAR-REGRESSION ESTIMATE

Graphical Representation of Linear-Regression Estimate



The DIFFERENCE for the CURRENT time frame
depends on the regression coefficient
based on previous survey results

The affect of this bias on the accuracy of an estimate made from the data depends on the ratio of magnitude of the bias to the magnitude of the standard error (the standard deviation of the sample divided by the square root of n-1). What happens, is that the actual area that remains under the probability sampling distribution curve (associated with the alpha level chosen) becomes distorted (asymetrical) and the actual chance of making an "alpha error" changes from that intended.

The following table indicates the impact that differing ratios of bias error to standard error have on the actual size of the upper and lower tails of a sampling distribution for an assumed unbiased alpha of 5 %.

AFFECTS OF A RANDOM SAMPLING ERROR

If the measuring method selected has a random error (e.g., plus or minus 1/10th centimeter, on the average) associated with it, (or if something else about the sampling procedure causes a randomly distributed error to be introduced into the sample data) then estimates made from the

TABLE 10 -- AFFECTS OF SAMPLING BIAS AT A 5 % ALPHA

$b/S_{\bar{x}}$	Actual Alpha	Actual Size Lower Tail	Upper Tail
0.00	0.0500	0.0250	0.0250
0.02	0.0500	0.0238	0.0262
0.04	0.0502	0.0228	0.0274
0.06	0.0504	0.0217	0.0287
0.08	0.0508	0.0207	0.0301
0.10	0.0511	0.0197	0.0314
0.20	0.0546	0.0154	0.0392
0.40	0.0685	0.0091	0.0594
0.60	0.0921	0.0052	0.0869
0.80	0.1259	0.0029	0.1230
1.00	0.1700	0.0015	0.1685
1.50	0.3231	0.0013	0.3228

sample data will project a correct estimate for population values; but the standard error of the estimate will be increased by the standard error of the measurement error. That is, if one can determine the actual errors that were made during the data taking, were to calculate their standard deviation, and then divide it by the sample size; then this standard error of the measurement error should be added to the standard error of the data to determine the actual standard error of the sample.

$$S_{\bar{x}\,\text{Actual}} = S_{\bar{x}\,\text{Calc}} + S_{\bar{e}\,\text{Error}}$$

where:

$$S_{\bar{e}} = S_{e} * [\ 1/n^{**}(1/2)\ ]$$

PROBLEM SOLVING TECHNIQUES

Forecasting Techniques

SAMPLING FROM FINITE POPULATIONS

In general, sampling from finite populations is mathematically more complex than sampling from infinite populations (or than sampling with replacement from finite populations). Fortunately, the magnitude of the error remains small (negligible) until the sample comprises greater than 5 to 10 % of the total population. Beyond this, correction factors need to be applied to give credit for the "extra" informational content that has been obtained.

For example, a common correction factor for the standard error is:

(N-n/N-1)**(1/2)

Where:

N = Population size

n = Sample Size

This correction factor goes to 1.0 (i.e., has no impact) when the sample is small compared to the total population size.

INFERENTIAL STATISTICS

SOME DEFINITIONS

Hypothesis

A hypothesis is a statement of an expected relationship between variables that may be tested empirically to determine its validity. A hypothesis may be derived from observation, deduced from a larger body of theory, or based simply on a hunch that the analyst is willing to use provisionally.

Hypothesis Testing

Hypothesis testing refers to the attempt to confirm or disconfirm a factual proposition, or hypothesis, by gathering and analyzing relevant evidence. Testing refers both to the use of certain types of measuring instruments, such as intelligence tests or attitude scales, and to statistical tests of the level at which observed relationships in data are sufficiently different from chance relationships (level of significance). In the latter case, a null hypothesis, or opposite of the original hypothesis, is formed with the expectation that it can be disproved. If the null hypothesis can be rejected, confidence in the original hypothesis is increased.

Null Hypothesis

The null hypothesis is the opposite of an original hypothesis, one which an analyst formulates with the expectation that it will be discredited by the data gathered. One using the scientific approach does not usually try to prove directly that a hypothesis is true but seeks to discredit its opposite with a given degree of probability. For example, it may be hard to prove directly that a certain strain of virus causes a disease. The theory may be supported, however, if diet, weather, insects, other viruses, etc., can be discarded from consideration. If the test of the null hypothesis shows it can be rejected with 99% confidence, the original hypothesis is strengthened, but not proven true.

Alternative Hypothesis

An alternative hypothesis, in the testing of hypothesis, is any admissible hypothesis alternative to the one under test (i.e., alternative to the Null Hypothesis).

PROBLEM SOLVING TECHNIQUES

Forecasting Techniques

Significance Test

A significance test is a statistical test designed to determine the probability that the distribution of observed data could occur entirely by chance. These tests generally specify some low probability level, such as 1 or 5 percent, as the threshold for consideration of the data. This means that unless we are 99 or 95 percent certain that the observed distribution could not occur by chance, the data is considered insignificant. Some of the more common tests of significance include: chi square, F test, Fisher exact test, Kolmogorv-Smirnov D test, t test, and z test.

Confidence Level

The confidence level refers to the statistical probability derived from a significance test (such as chi square). It tells how often the observed behavior (if the null hypothesis were true) would occur solely due to chance. It, therefore, gives information about how certain we may be that non-random causes are operating.

Significance Level

The significance level is a term equivalent to the confidence level.

Acceptance Region

The acceptance region is one of two mutually exclusive regions in a sample space used in the testing of statistical hypotheses. If the sample point falls into one (the region of acceptance) the hypothesis cannot be rejected; if it falls into the other, the hypothesis is rejected.

Critical Region

The critical region is the other of the two mutually exclusive regions in a sample space used in the testing of statistical hypotheses. If the sample point falls into the critical region, the hypothesis is rejected.

Alpha Error

An alpha error in hypothesis testing is the rejection of a hypothesis when it is actually true.

Beta Error

A beta error is accepting (failing to reject) a hypothesis when it is actually false.

Matched Samples

Matched samples refer to a pair or set of samples in which each member of a sample is matched with a corresponding member in every other sample by reference to qualities other than those immediately under investigation. The object of matching samples is to obtain better estimates of differences by removing the possible effects of extraneous variables. A survey of political attitudes, for example, might match individuals of several different samples according to age, income, education, race, and sex, but not residence, in order to ascertain regional differences in political beliefs.

CONFIDENCE INTERVALS

Interval Data

Confidence intervals for interval data are calculated by using the sample mean and standard deviation to determine a range of values that should contain, at a specified probability level, the true value of the population mean. This is done by using the sample mean as a best-guess estimator of the population mean, and then using the standard deviation to estimate the probable range of error on either side of the sample mean.

It turns out that the sample mean becomes a better and better estimator of the true population mean as the sample size increases. The mathematical relation is that the size of the probable error of using the sample mean for the population mean decreases as one over the square root off the sample size. It also turns out that the probability density function of the actual misses that would occur for a large number of samples is normally distributed around the sample mean with a standard deviation (usually referred to as the standard error) equal to the sample standard deviation divided by the square root of the sample size.

$$s_{error} = s/(n^{**}(1/2))$$

By using statistical tables (such as the Normal or Z table, or the Student's t table; depending on such things as sample size) one can determine confidence intervals that exclude (leave to chance--or more precisely, alpha error) as small a tail of the probability density function curve as is desired. Of course, the smaller the chosen alpha error, the larger (less definite) the calculated confidence interval.

Ordinal Data

Confidence intervals for ordinal data are determined via exact probability calculations. The actual values of the sample data are ignored (except for use in determining their ordinal ranking). Exact probability calculations are then made to determine the chance of selecting, randomly, a certain number or fewer of the sample members all from above (or below) the median.

The calculations, which can get tedious, do not, in general, need to be performed each time a confidence interval is to determined. Values for common sample sizes and common alpha levels can be obtained from tables found in most statistics texts. The tables do not give the confidence limit values directly, but rather indicate the ranks of the smallest and largest sample observation that represent the extreme limits of the confidence interval.

HYPOTHESIS TESTING OVERVIEW

In the section on Confidence Intervals we saw that levels of certainty can be calculated on sample means, variances (standard deviations), and medians. These confidence levels tell us how likely it is that the true population parameter value that we are approximating lies outside of the interval we determined. We could instead make some hypothesis about a population parameter, calculate an interval that should include a stipulated percentage of sample estimators of the parameter, and then take a random sample from the population to see if we obtain an estimator value that falls into the "expected" range. By proceeding in this fashion, we can make a quantitative probability statement about correctness of our initial hypothesis for the population parameter. This general procedure forms the basis for all of the hypothesis tests that we shall consider in BSCAM. (The "Confidence Intervals" may seem obscure tc you on several of the specific tests we cover, but the principle still applies. That is, we will simply be choosing some estimator for which we can theoretically calculate a probability distribution, and will then determine whether the estimator we calculate on a random sample falls into an expected range of values, or if it instead lies far out in one of the tails of the distribution.)

To perform hypothesis tests correctly we need to be familiar with a few logical principles, some procedural guidelines, and the specific calculations that are associated with each statistical test. The remainder of this section will ccver these three things.

Logical Principles

It is virtually impossible to prove something to be unquestionably true. Some untested "what if" scenario can almost always be devised. However, the occurrence of a single contradiction can prove a hypothesis to be false--i.e., to be infeasible. Hence, hypothesis tests are designed to disprove a hypothesis (the <u>null hypothesis</u>). To do this effectively we must make two hypotheses, the null hypothesis, and an <u>alternative hypothesis</u>. To be useful, the two hypotheses should be both <u>mutually exclusive</u> and <u>exhaustive</u>. That is, at least one of the two must be true, but both of them can not be true--i.e., exactly one of the two must be true. For example:

Null Hypothesis (HO): $\bar{X} \geq$ "value"

Alternative Hypothesis (H1): X < "value"

Our goal is not to disprove the null hypothesis completely. Instead, we will use hypothesis tests to determine if our sample data is less likely than to have been obtained than some predetermined likelihood criteria. That is, we shall choose some probability level (a so-called <u>alpha level</u>), and will reject our null hypothesis if our sample is less likely to have been obtained than our alpha level if the null hypothesis were true. This does not prove the null hypothesis wrong, but does give us (1-alpha)*100% confidence that we have not made a mistake by rejecting H0.

Since H0 and H1 are mutually exclusive and exhaustive, if we reject H0, then we must assume H1 to be true. If our sample is not less likely than the designated alpha level, then we cannot reject the null hypothesis. This does not show H0 to be true, however. It rather means that one of the following situations should be assumed to be true:

a) H0 is true,

b) H0 is not true, but "things" have not changed sufficiently for us to detect the change. (That is, the change has only been a small--and probably unimportant--one.)

c) H0 is not true and things have in fact changed substantially, but Mother Nature fooled us by providing a very unlikely sample that did not evidence the change.

From the discussion of the preceding paragraphs it should be evident that two types of errors can be made when making a decision based on a hypothesis test. These two types of errors are called Type I (or alpha) errors and Type II (or beta) errors. Table 11 illustrates these errors on the basis of a hypothesis test decision and the actual state of nature.

TABLE 11 -- TYPES OF INFERENCE ERRORS

		Actual State of Nature	
		Ho is True	Ho is False
Decision	Reject Ho	Type I Error	Correct Decision
	Accept Ho	Correct Decision	Type II Error

The type I (alpha) errors correspond directly to the certainty levels we discussed above in deciding whether or not to reject H0. The Type II (beta) error corresponding to the b) outcome two paragraphs ago is more difficult to predict. We shall worry primarily about Type I errors, but don't forget that we may be making Type II errors as we proceed. Back in the real world we must note in each case whether a beta error would be a serious error.

Performing statistical tests on data requires certain procedures to be followed if unbiased and meaningful test results are to be obtained. For instance, the appropriate test to use and the level of "unusualness" of the data required to reject the null hypothesis (the alpha level) should be determined prior to collecting any data. That is, one should not collect some data and then look for some test that will support preconceived ideas about the data, nor should one test the data and then establish an acceptable alpha level evaluating the obtained alpha level of the test statistic.

The cookbook-type protocol shown in Table 12 provides guidance for avoiding most of the common mistakes in applying statistical tests to data.

TABLE 12 -- STEPS IN HYPOTHESIS TESTING

1. Examine the type of data needed or expected. (Interval, Ordinal, etc.) (Population parameters known?) (Random?) (Distribution shape?)

2. Define the Null Hypothesis and Alternative Hypotheses.

3. Choose the appropriate test statistic. (t, Z, F, etc.)

4. Set the level of significance (the alpha level) that you will require to reject the null hypothesis and, if possible, decide what sample size is required.

5. Convert the alpha level to the critical region of the test statistic.

6. Collect your data.

7. Evaluate the test statistic and use it either to reject the null hypothesis or to not reject it.

Overview of Hypothesis Test Types

ANALYSIS ACTIVITIES

Specific hypothesis tests can be thought of in a number of different ways. For example, you may like to remember tests according to their theoretical basis (tests having a common theoretical derivation). For our purpose here, an alternative approach seems more appropriate. We'll group tests into categories based on their purpose or use (e.g., tests designed to check for changes in population location [means, median, etc.] or to measure dispersion [e.g.,variance])

The following table cycles through tests that apply to given data types, this first for tests comparing two data sets at a time, and then for tests designed to handle multiple data sets simultaneously (ANOVA tests).

For each data type (interval, ordinal, and binomial) several tests are listed, each with a slightly different use or purpose. The table also lists goodness-of-fit tests, that measure how a set of data is distributed. (Some of these test also can be useful for testing nominal data.

TABLE 13 -- TYPES OF HYPOTHESIS TESTS

INTERVAL SCALE DATA

- Difference in Means
  - Sample versus Population
    - Large n
    - Small n
  - Sample versus Sample
    - Independent Samples
    - Matched Samples
- Difference in Variance
  - Sample versus Population
  - Sample versus Sample

ORDINAL SCALE DATA

- Difference in Medians
  - Sample versus Population
  - Sample versus Sample

BINOMIAL DATA (PROPORTIONS)

- Sample versus Population
- Sample versus Sample

GOODNESS-OF-FIT TESTS

- Nominal Data
- Interval (or Ordinal) Data

ANALYSIS OF VARIANCE (ANOVA)

- Interval Scale Data
  - One-Way ANOVA
  - Two-Way ANOVA
- Ordinal Scale Data
  - One-Way ANOVA
  - Two-Way ANOVA

PROBABILITY FORECASTING

This section has been adapted from a portion of the HANDBOOK OF FORECASTING TECHNIQUES, Part II, Description of 31 Techniques, prepared for the U.S. Army Engineer Institute for Water Resources, Fort Belvoir, VA. The work was done under Contract No. DACW31-75-C-0077, August 1977.

ABSTRACT

Probabilistic forecasting encompasses a number of different techniques all of which use some aspect of probability theory. Its major use is to predict the probability of occurrence of future events for some time dependent random process. Some of the probabilistic forecasting techniques are:

Point and interval estimation

Monte-Carlo simulation

Markov processes

Parametric sensitivity analysis

Inventory theory

Queueing theory

Generally, all of these techniques require modeling of events and estimations of probability functions.

ANALYSIS ACTIVITIES

Forecasting Techniques

DEFINITION

Probabilistic forecasting encompasses a number of different techniques all of which use some aspect of probability theory. Probability theory is regarded here as the study of mathematical models of random phenomena. A random phenomenon is defined as an empirical phenomenon that obeys probabilistic, rather than deterministic laws.

A random phenomenon that is time developing in a manner controlled by probabilistic laws is called a stochastic process. Thus the motion of a particle in Brownian motion, the growth of a population such as a bacteria colony, and the occurrence of floods on a river are examples of stochastic processes. In general, probabilistic forecasting is the result of a stochastic process.

MAIN USES

The major use of probabilistic forecasting is to predict the probability of occurrence of future events for some stochastic process. Changes or modifications to the stochastic process can be evaluated by the use of parametric sensitivity analysis. For example, the flooding of a river can be evaluated as a stochastic process. The introduction of various flood control measures on the river can be included in the evaluation by means of parametric analysis.

LIMITS AND CAUTIONS

Probabilistic forecasting is based on the development of a stochastic model to duplicate the hypothetical occurrence of events in the real world. The validity of the results of the forecasts are, therefore, dependent on how well the stochastic model duplicates the real world and on the availability and precision of inputs required by the model. The general procedure used to overcome this difficulty is to "bracket" the outcomes hypothetical forecasts by varying inputs over their range of uncertainty by the use of parametric sensitivity analysis.

OTHER TECHNIQUES

There are a number of alternatives available to probabilities forecasting. Historically, subjective opinion by "experts" has been the most used alternative. The use of subjective opinion by "experts" has been structured in recent years into methodologies of their own (e.g., the Delphi technique). Deterministic modeling is another alternative that is generally applied to very complex systems. The results of the deterministic models are usually labeled as experted values with the term, experted, being used in a rather loose statistical sense.

The entire field of decision analysis can be regarded as a subset of probabilistic forecasting; however, it is treated in more detail in the section on decision techniques--starting on page 237.

PROCEDURES

Probabilistic forecasting uses a number of different techniques. There are some steps, however, that are common to all of the techniques. These are:

Detailing all possible events.

Estimating probabilities of occurrence of events subject to various conditions.

A modeling effort which is specific to the forecasting technique to be applied.

Descriptions of some of the probabilistic forecasting techniques that may be used singularly or in combination include:

Point and interval estimation.
This method (which is discussed in detail starting on page 153) attempts to describe the probability of outcomes for a single event. An example would be the maximum annual water level of a river. The event of concern would be a flood, i.e., when the water level exceeds a given depth.

Monte-Carlo Simulation.
This method simulates all or part of a process by running a sequence of events repeatedly, with random combinations of values, until sufficient statistical material is accumulated to determine the probability distribution of the outcome. The values used in each run are selected because of the assumed underlying statistical distribution.

Markov Process.

The procedure used in the Markov process is to model the situation based on a matrix of transition probabilities. These transition probabilities are the probability of going from one event to another and are assumed to be independent of the sequence of events. Points of interest to the Markov process are the expected times required to go from one event to another and the steady state solutions, i.e., the long-run probability of being in any state.

Non-Markov Process.

This process is similar to the Markov process, only it is assumed that the transition probabilities can depend on the preceding sequence of events.

Parametric Sensitivity Analysis.

In this type of analysis (also discussed in the section beginning on page 168), the input parameters or model formulation are varied in a systematic manner to evaluate their effect on possible outcomes. Variables whose variation have little or no effect on possible outcomes can then be treated in a deterministic manner.

Inventory Theory.

Two problems of considerable importance to retail shops, wholesale distributors, manufacturers, and consumers holding stocks of spare parts are: (1) deciding when to place an order for replacement of their stock items, and (2) deciding how large an order to place. Two kinds of uncertainty must be taken into account: (1) number of items that will be demanded during a time period, and (2) time-of-delivery lag that will elapse between order and receipt of goods. Inventory control is concerned with minimizing the cost of maintaining inventories, while at the same time keeping a sufficient stock on hand to meet all contingencies.

Queues.

A queue (or waiting line) is generated when customers (or servers) arriving at some point to receive (or render) service there must wait to receive (or render) service. In the mathematical theory of queues, waiting lines are classified according to four aspects: (1) the input distribution (the probability law of the times between successive arrivals of customers); (2) the service time distribution (the probability law of the time it takes to serve a customer); (3) the number of service channels; and (4) the queue discipline (the manner in which customers are selected to be served; possible policies are "first come, first served," random selection for service, and service according to order of pricrity). Queueing theory is concerned with the effect that each of these four aspects has on various quantities of interest, such as the length of the queue and the waiting time of a customer for service.

PRODUCT OR RESULTS

The actual results depend on the procedure employed. Generally, the results include tabular and graphical presentations. These might give expected outcomes over time and their associated probability intervals. If parametric sensitivity analysis is used, there might be a different set of results for the conditions studied.

COMMUNICABILITY OF RESULTS

Some knowledge of basic statistics is required to understand the results. Otherwise, the results, which are presented in tabular and graphical form, can be readily comprehended. Because of the nature of the modeling and estimation of probabilities required, the results can be subject to criticism. This, however, is less of a problem than with other forecasting techniques.

SPAN OF FORECASTS

The most accurate forecasts are for the near and medium term time frames. As the time span increases, the validity of the underlying assumptions decreases. Moreover, it is more difficult for the person or persons developing the model to enumerate all possible events for the longer-term situations.

RESOURCES

The resources required to conduct probabilistic forecasting include experts in the applied subject area and in statistics who are able to model events, data, and estimates of conditional probability distributions for the occurrence of events, given the occurrence of other events; and many of the forecasting techniques require the use of a computer. The cost of te development of probabilistic forecasting models and their associated inputs may require anywhere from a few dollars to many thousands of dollars, depending on their size, scope, and application.

ANALYSIS ACTIVITIES

SPECIAL COMMENTS

Since there are a number of different probabilistic forecasting techniques that can be used with varying levels of detail, the validity of the forecast will depend on the model and data used. A carefully designed model used with a good data base and using parametric sensitivity analysis to explore the effect of less reliable inputs can give extremely useful results.

PROBLEM SOLVING TECHNIQUES

Forecasting Techniques

CORRELATION AND REGRESSION

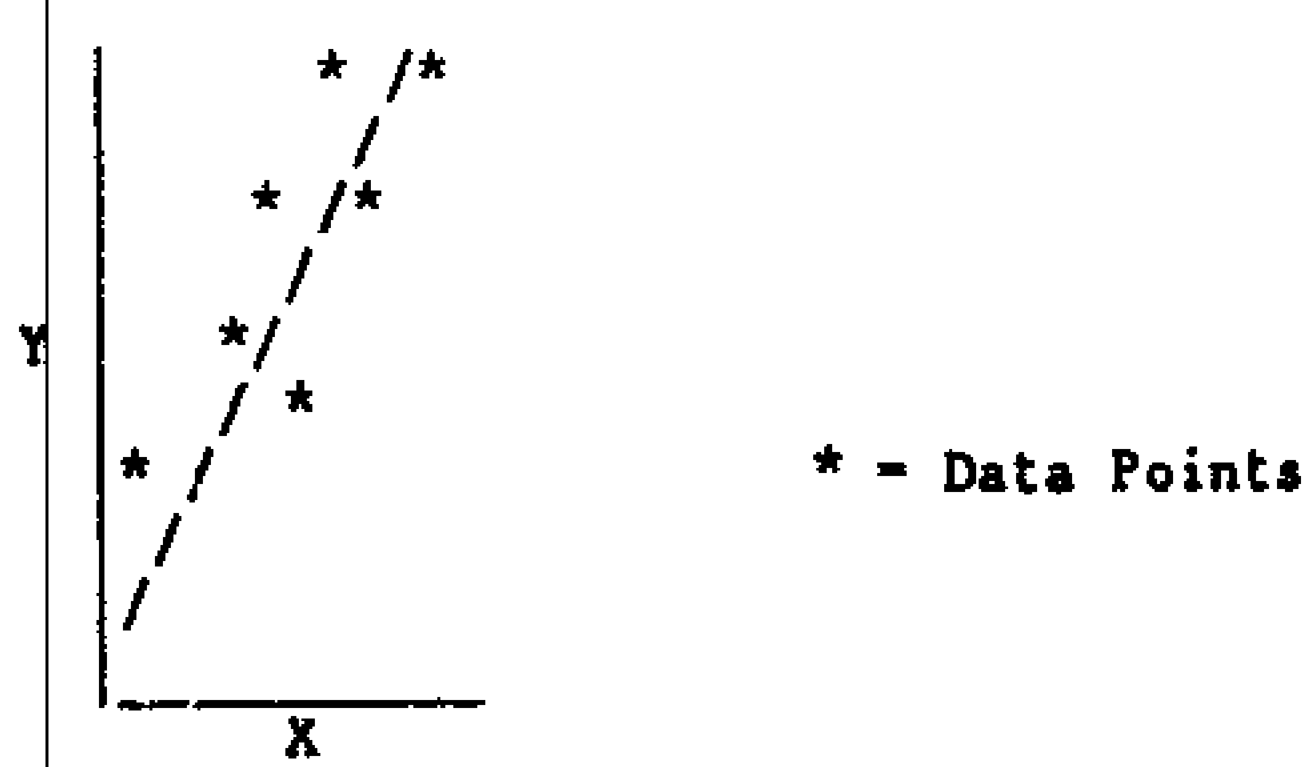
TERMINOLOGY

Correlation

A measure of the degree of association between two or more sets of data. That is, a measure of the degree to which two variables are related.

Regression

A technique for predicting the value of some unknown parameter based only on information about the current values of other parameters. Regression makes use of both the degree of association that exists among variables (determined via correlation techniques) and the mathematical function (usually a straight line) determined (usually by the method of least squares calculations) to best describe the relationships present among variables. If values from only one independent variable are used to predict values for another, dependent variable, then the process is referred to as bivariate regression. Multivariate regression involves using values from more than one independent variable to predict values for a dependent variable.



Least Squares Method

The least squares method is a mathematical technique that specifies the regression line which represents the "best fit" possible of a set of data. The resulting regression line is calculated in a manner that minimizes the sum of the squares of the vertical "miss" distances of the data points and the determined line of regression.

Correlation Coefficient (Calculated or Experimental)

Both r and its square (r^2) are obtained during correlation calculations. They provide a measure of the degree of association between the dependent variable and an independent variable. The percent of the variation of the dependent variable that can be explained or determined by associated changes in an independent variable is given by r^2. Other names frequently used for r and r^2 are the Coefficient of Determination for r^2 and the Correlation Coefficient for r. When used to describe the overall information about variations that occur in the dependent variable based on simultaneous consideration of more than one independent variable, then r may be referred to as the Multiple Correlation Coefficient. The formula for computing the calculated (experimental) value of the correlation coefficient is:[29]

$$r = \frac{n \bullet (\$XY) - \$X \bullet \$Y}{[(n(\$X^2) - (\$X)^2) \bullet (n \bullet \$(Y^2) - (\$Y)^2)] ** (1/2)}$$

Correlation Coefficient (Theoretical)

If one wants to be assured (to some specified alfa level) that the correlation coefficient calculated on the data is not do to chance, a theoretical correlation coefficient can be computed that gives the the smallest "r" that would be found by chance at the chosen alfa level of confidence. The formula for computing the theoretical correlation coefficient is:

$$r_{.05} = \frac{t^2}{t^2 + (n-2)}$$

(where n is the number of observations of the independent variable)

---

[29] The symbol \$ will be used in this section to represent the summation sign -- normally represented by a capital Greek sigma, a character not included in the print train used for this book.

PROBLEM SOLVING TECHNIQUES

Forecasting Techniques

EXAMPLE PROBLEM

	10 Billions of KWHR	Metric Tons of Iron			
	X	Y	X•Y	X^2	Y^2
	54	30	1620	2916	900
	60	33	1980	3600	1089
	66	37	2442	4356	1369
	71	40	2840	5041	1600
	76	46	3496	5776	2116
	81	50	4050	6561	2500
$	408	236	16428	28250	9574
/n	68	39.33			
•n			98568	169500	57444

CALCULATION OF THEORETICAL r

$$r^2_{.005} = \frac{t^2}{t^2_{.005} + (n-2)} = \frac{(5.959)^2}{(5.959)^2 + (6-2)} = \frac{35.51}{39.51} = 0.8988$$

$r_{.005} = 0.9480$

If the Computed r is equal to or greater than this Theoretical r, all is well. (i.e., there are only 5 chances in 1000 of getting a calculated correlation coefficient as large as 0.8988 if the are in fact not correlated.)

CALCULATING THE COMPUTED r

$$r = \frac{n \bullet (\S XY) - \S X \bullet \S Y}{[(n(\S X^2) - (\S X)^2) \bullet (n \bullet \S(Y^2) - (\S Y)^2)] ** (1/2)}$$

$$= \frac{98568 - (408) \bullet (236)}{[(169500 - 408^2) \bullet (57444 - 236^2)] ** (1/2)}$$

$$= \frac{98568 - 96288}{[(169500-166464) \bullet (57444 - 55696)] ** (1/2)}$$

$$= \frac{2280}{[(3036) \bullet (1748)] ** (1/2)} = \frac{2280}{2303.68} = 0.9897$$

Since the Computed r is greater than the Theoretical r we may continue.

CALCULATION OF SLOPE for the REGRESSION LINE

$$m = \frac{n\S(X \bullet Y) - \S X \bullet \S Y}{n \bullet \S X^2 - (\S X)^2} = \frac{98568 - 96288}{169500 - 166464}$$

$$= \frac{2280}{3036} = 0.751$$

CALCULATION OF INTERCEPT for the REGRESSION LINE

$$b = \overline{Y} - m\,\overline{X}$$

$$= 39.33 - (0.751)(68) = 39.33 - 51.07 = -11.74$$

PROBLEM SOLVING TECHNIQUES

LINE OF REGRESSION

$Y = m X + b$ $Y = 0.751 X - 11.74$

Sample Projections

for $X = 85$

$$Y = (0.751)(85) - 11.74 = 52.10$$

for $X = 90$

$$Y = (0.751)(90) - 11.74 = 55.85$$

BACKGROUND REFERENCES

Korin, Basil P., Statistical Concepts for the Social Sciences, Winthrop Publishers, Inc., Cambridge, Massachusetts, 1975.
95-133
303-321

Remington, R.D. and M.A. Schork, Statistics with Applications to the Biological and Health Sciences, Prentice Hall, Inc., Englewood Cliffs, New Jersey, 1970.
253-278

OTHER STATISTICAL TECHNIQUES

A forecast is a statement of what is expected to happen. It is based on past data and requires more than a single observation. Forecasting credibility depends on two factors proven true by future data. There are several numeric techniques which we will look at in order of increasing credibility:

Averages

Moving Averages

Weighted Moving Averages

Exponential Smoothing

Auto Correlative Techniques

Time Series Analysis

Adaptive Techniques

AVERAGES

Are measures of central tendency and fail to reflect current dynamic behavior.

MOVING AVERAGES

Attempt to solve this states condition by establishing a "data window" of arbitrary size to look only at the most recent in events. They treat all observations as equally important in their contribution to the future.

WEIGHTED MOVING AVERAGES

Assign a multiplier to each observation to give current observations more "weight" in the prediction. Normally all the weighting factors sum up to one. It is a compromise between responsiveness and accuracy.

EXPONENTIAL SMOOTHING

Based on past predictions of the historical time series as modified by current data. The past predictions provide a stable base line which, at the forecaster's option, may be modified to reflect deviations in a rapid dynamic manner (responsiveness) or in a cautious manner (accuracy). For oscillating data this causes a perpetual overshoot each time the observations change the sign of their slope.

AUTO CORRELATION

Identifies the periodicity and amplitude of sineusoedal data where there is a regular periodicity to the observational behavior and works well with a single type of repeated oscillation.

TIME SERIES ANALYSIS

Decomposes the data into a trend line, a seasonal variation, a critical variation and random noise and treat each individually but fails to react to change.

ADAPTIVE TECHNIQUES (BOX-JENKINS)

Adaptive techniques are a composite of the foregoing. They use coefficients which change with the impact of current observations so that each forecasted observation is based on newly recalculated parameters/ coefficients which are updated dynamically as new observations become available. The Box-Jenkins is one such technique.

This section has been adapted from a portion of the HANDBOOK OF FORECASTING TECHNIQUES, Part II, Description of 31 Techniques, prepared for the U.S. Army Engineer Institute for Water Resources, Fort Belvoir, VA . The work was done under Contract No. DACW31-75-C-0077, August 1977.

ANALYSIS ACTIVITIES

Forecasting Techniques

Abstract

The Box-Jenkins Method is a powerful forecasting tool useful where the variable of interest is a complex function of a variety of factors. It illuminates the pattern of the time-series data and uses that pattern to generate the forecast. It does this through multiple iterations gradually refining the model. As a result, however, it is complex to apply, costly, time consuming and requires extensive computer time.

Definition

The Box-Jenkins method is a quantitative forecasting approach designed to handle complex time series data where no ready pattern to the data is apparent. Unlike most other quantitative techniques Box-Jenkins does not require a clear definition of the trend. In fact, the method is intended to discover the pattern of the trend. This is accomplished by multiple iterations of the data, with each iteration providing information which allows the next to more closely approximate the actual pattern of the data.

History

It was developed in the late sixties by Professors G.E. Box and G.M. Jenkins and published first in 1970 in their book Time Series Analysis.

Main Uses

This method is useful where the desired forecast has an extended time span, the data composing the time series is complex and where a relatively high degree of precision is both required and meaningful. It is also useful in that the method itself provides information on the statistical error in the forecast independent of other statistical tests.

Limits and Cautions

Like all quantitative forecasting techniques relying on historical data, the forecast at best can only be as good as the data used. Thus, the quality of the data is an important and often ignored limit in forecasting. The Box-Jenkins method in particular is one of the most complex and costly methods of quantitative forecasting. It requires considerable computer time and generally longer time to complete a forecast. Its complexity demands a high degree of skill in mathematical methods.

Other Techniques

Box-Jenkins sits somewhere between multiple regression analysis and computer simulation models. It is a good deal more complex than multiple regression. However, it forecasts only a single variable and is therefore less powerful than simulation models.

Procedures

Initially, the person using the Box-Jenkins method postulates a general class of forecasting methods for his particular situation. In stage 1 a specific model that can be tentatively entertained as the forecasting method best suited to that situation is identified. Stage 2 then consists of fitting that model to the available historical data and then running a check to determine whether it is adequate. If it is not, the approach returns to stage 1 and an alternative model is identified. When an adequate model has been isolated, stage 3 or 4 is pursued stage 3 being developed of a forecast for some future time period and stage 4, the development of a control algorithm for a situation in which the forecasting method is to be used for control purposes.

Product or Result

Box-Jenkins produces two results. The first is the forecast itself in either tabular opr graphical form. It provides the most likely value as an upper and lower bound to the probability range. Equally important is the identification of a model for the data which may be used to generate future forecasts.

5

Level of Confidence of Results

The method includes as part of the approach, statistical techniques for evaluating the confidence. In general, however, Box-Jenkins tends to be more accurate than the less complex methods.

Span of Forecasts

Box-Jenkins is suitable for forecasts of the very short-range to the intermediate level of a few years. Beyond that the reliability of the identified pattern derived from the complex of variables cannot be counted on.

Resources

The necessary resources to successfully apply the Box-Jenkins method are considerable. It is costly and time consuming. It requires extensive computer time. A high level of mathematical skills is also necessary.

PROBABILISTIC REASONING

AN INTRODUCTION TO PROBABILISTIC THINKING

INTRODUCTION

Preliminary results of research being conducted by the Analyst Training Branch of the Office of Training and Education indicates that the majority of NFAC analysts have similar psychological profiles. The profile that dominates describes persons who tend to make early judgments about "THE correct answer" to a problem, and who then defend this initial judgment quite tenaciously. This tendency leads (usually without conscious awareness of the process) to the selection of data that support the favored hypothesis, and to the rejection (or dismissal as unimportant) of data that conflict with it. This approach breaks several tenants of good problem solving procedures: in particular, the tenant of postponing evaluation and judgment until all data have been made available.[30]

The primary danger of reaching one's conclusions too early (premature foreclosure) is not so much in making a bad assessment because the data are incomplete. Rather, the danger is that in situations where the world is changing quickly, or in situations where a major, unprecedented event occurs, one becomes trapped by the judgments already made and tends to increase chances of missing indications of change, and hence reduces chances of revising (changing) the initial estimate when this should be done.

Fortunately, there are several problem solving approaches that help to prevent premature foreclosure. They can be both simple and relevant to use in intelligence analysis. One of the more useful is probabilistic reasoning, for it can be used at any of several levels of complexity (each successive level of sophistication adding new capability and completeness). But even the simplest level helps to prevent premature foreclosure and serves to add structure to complicated problems.

----

[30] For a good general article that discusses the problems humans tend to encounter when making analytical judgments, see Biases in Evaluating Information, by Dick Heuer, initially published in ANALYTICAL METHODS REVIEW, A Review of Ideas and Applications, January, 1981; and included as a chapter in this Handbook.

PROBLEM SOLVING TECHNIQUES

Probabilistic Reasoning

GENERATING ALTERNATIVES

The first step to probabilistic reasoning is no more complicated than stating formally that more than one outcome is possible. This can be accomplished simply by listing all possible outcomes to the problem under consideration.

To be most useful (and to permit later extension to more sophisticated analyses) the alternatives generated should be mutually exclusive (only one can occur, not two or more simultaneously) and exhaustive (nothing else can happen, one of the alternates listed has to occur). For instance, if one wants photographic information about some geographic location, there are several possible outcomes to a request to take a picture. One could list two outcomes only:

1. Picture obtained
2. Picture not obtained

This list follows the mutually exclusive and exhaustive principles mentioned above. If a third option called "Picture obtained within 24 hours" was added, the mutually exclusive principle would have been violated (unless the first outcome had been reworded to "Picture obtained after 24 hours").

The above list may or may not be very useful with just two outcomes. If one is interested in the reason why a picture was not obtained, then the second outcome can (and should) be decomposed further. A revised list might look like:

1. Picture Obtained
2. Picture Taking Request Denied
3. Bad Weather Prevented Picture from Being Taken
4. Camera Malfunctioned
5. Film Was Damaged or Misdeveloped
6. Camera Was Lost
7. Film Was Lost
8. Picture Was Lost
9. Picture of Wrong Location Was Received

This list may well be too detailed, but illustrates the way that specifying all possible outcomes that are of interest can expand ones perspective about possibilities. The expanded possibilities often can generate useful insights to problems. For example, outcome 8 in the above list suggests that a phone call to see if the picture was taken might be useful if you don't receive a copy of "your" picture in the mail, since it may have gotten lost in the mail or someone may simply have forgotten to mail it to you.

The key is listing all of the outcomes that are meaningful to you. The list then can serve both as a reminder that multiple outcomes can occur, and as a checklist to decide how any item of new intelligence might affect your assessment of the relative likelihoods of the diverse outcomes listed.

ANALYSIS ACTIVITIES

Although there is no need to maintain every possible outcome on your working list, it usually is a good practice to make an initial list relatively detailed. It is far easier to combine multiple outcomes that you do not need to distinguish among, than it is to think of something new that wasn't listed or to think of separating one combined-event outcome into its subcomponents. The mere act of generating a complete, detailed list often provides a useful perspective on a problem.

When generating a set of outcomes, beware of using generic terms (such as "other"), since humans usually are poor at recalling the frequently great extent of things that could fall under that seemingly simply lable. This does not mean that a catch-all outcome label should not be used, but rather that it should be included only when a complete list of all alternatives cannot be generated first. Also, do not overlook the outcome possibility of nothing happening. For instance, if one is trying to generate a list of all the things that the Kremlin might do regarding some issue, one of the often neglected possibilities is to do nothing at all.

DESCRIBING INFLUENCES

The list of alternatives generated above represents the possible outcomes of an "event." By thinking in terms of the event, we can ask ourselves which of the listed outcomes will be the one to occur. But, in answering this question we may be tempted to respond "It all depends." This is common. What we will do now is develop a shorthand technique (that starts where the alternative-listing technique left off) for indicating the things on which the outcome selection depends, and for indicating the interrelationship of the influences we enumerate.

For instance, let's assume that we want to assess the outcome of an upcoming election in a Latin American country. There are three candidates: one (the incumbent) is pro-West, one is pro-Soviet, and the third is Isolationist. We know that the election will be held and that one (and only one) of the candidates will win. From this information we can generate a list of alternative outcomes, just as we did earlier, that looks like this.

1. Pro-West
2. Pro-Soviet
3. Isolationist

To distinguish these alternative outcomes as belonging to one common event, we simply add the event name (ELECTION RESULT) above the list of alternative outcomes.

ELECTION RESULT
1. Pro-West
2. Pro-Soviet
3. Isolationist

We now are ready to describe those things that will influence our assessment of the relative likelihoods of the outcomes listed under ELECTION RESULTS. This is done by listing all of the "uncertain events" that will influence the relative likelihood of the ELECTION RESULT Outcomes, and then drawing an arrow from the name of the influencing event(s) to the "ELECTION RESULT" block. Uncertain events are ones that, like the event of main interest, have several possible outcomes. For instance, ELECTION RESULT could depend on whether the COST OF LIVING Improves, Remains the Same, or Deteriorates in the time period just before the election. It also might depend on the success of the pending WHEAT HARVEST, that can be either a Bumper Crop, Adequate, or Inadequate. In addition to listing the names of these influencing events, we also will treat them just as we did the event of main interest, and list the alterative outcomes of each event by the event's name. (Figure 26 shows the structure, but you'll need to insert the arrows to complete the diagram.)

FIGURE 26 -- SAMPLE INFLUENCE DIAGRAM

ELECTION RESULT
1. Pro-West
2. Pro-Soviet
3. Isolationist

COST OF LIVING
1. Improves
2. Status Quo
3. Deteriorates

WHEAT HARVEST
1. Bumper Crop
2. Adequate
3. Inadequate

After we have added these influencing events we can use the diagram as an aid to prompt more information. Namely, we ask at least two questions of ourselves.

1. Do any of the influencing events we added influence each other?

2. Can we assess the relative likelihood of the outcomes of the newly added (bottom node) events directly, or do the outcomes of these events depend in turn on other influencing events (and outcomes)?

If the answer to the first question is that the added nodes influence each other, then arrows indicating the direction of influence should be added accordingly. In the above, assume that the WHEAT HARVEST Outcome will influence the COST OF LIVING Outcome, but not visa versa. We therefore add (using Figure 27) an additional arrow drawn from the WHEAT HARVEST node to the COST OF LIVING node.[31]

FIGURE 27 -- INFLUENCE DIAGRAM WITH PARALLEL INFLUENCE SEQUENCES

ELECTION RESULT
1. Pro-West
2. Pro-Soviet
3. Isolationist

COST OF LIVING
1. Improves
2. Status Quo
3. Deteriorates

WHEAT HARVEST
1. Bumper Crop
2. Adequate
3. Inadequate

The second question, having to do with additional influences not yet shown on the diagram, allows one to extend this pictorial representation of influences to whatever level of detail one feels is necessary. As is about to be discussed, however, beware of adding layers of detail that aren't needed. Making things more detailed <u>than is necessary</u> can degrade, rather than improve, the usefulness of this diagramming technique.

The process of going through this structuring exercise can be useful in several ways. For one, it helps identify <u>and document</u> those factors that are considered relevant to making judgments about the relative likelihoods of the main-event outcomes. The documentation (or audit trail, as it sometimes is called) feature is particularly useful in showing colleagues what your thinking has been, especially if you desire help in upgrading the diagram with things you may have overlooked. The process of structuring also helps in estimating the relative importance of the listed influencing events; since one can quickly determine from the diagram whether a given influencing factor affects your assessment of the main event likelihood

----

[31] The diagramming technique being discussed will <u>not</u> be impaired if the influence arrows comprising the diagram form a closed, unidirectional (feedback) loop; but the probability calculations and the tree structuring techniques that will be discussed later demand that feedback loops do <u>not</u> exist. If feedback loops emerge and are needed in your diagrams, you will need to use techniques designed to handle dynamic feedback situations--e.g., causal loop analysis or simulation modeling.

outcomes directly, or indirectly through one or more intermediate factors (and hence has a diluted impact).

Another advantage accrues from the thought process that is invoked when generating the list of influencing events and their outcomes. Namely, in the process of generating these lists one must address the issue of whether the event (or outcome) being listed actually will make a difference in your assessment of the relative likelihood of the outcomes of any of the events being listed. For instance, if in the WHEAT HARVEST example, we knew that an adequate harvest would insure the re-election of the incumbent, and that a bumper crop would do the same thing, then we have no need to differentiate these as two separate outcomes. We should instead simplify the diagram to that shown in Figure 28.

FIGURE 28 -- PRUNED INFLUENCE DIAGRAM

ELECTION RESULT
1. Pro-West
2. Pro-Soviet
3. Isolationist

COST OF LIVING
1. Improves
2. Status Quo
3. Deteriorates

WHEAT HARVEST
1. Adequate
2. Inadequate

Our thought process also should help us identify those events which contain no uncertainty. For instance, one might feel (quite correctly) that the age of the candidates in our example might affect their chances of being elected. However, their ages are <u>not</u> an event that will change in an unpredictable or uncertain manner. We can know their ages and can therefore specify precisely how we think this will affect the candidates chances of getting elected. We <u>make use of</u> this information when assessing the relative likelihoods of the main-event (ELECTION) outcomes, but we do <u>not</u> need to include it in our diagram of the influences of uncertain events.

EVALUATING OUTCOME SCENARIOS

The diagram for the example problem we have been developing could grow larger and more complicated, but to keep things somewhat simple for now, let's assume that our diagram is complete. This implies that we are comfortable with evaluating the relative likelihood of the options listed within each of the "bottom node" events; otherwise we would have added additional influence nodes to the diagram.

Inspection of the final diagram indicates that we have specified 18 different, mutually exclusive future scenarios.[32] To determine the 18 scenarios specified by the influence diagram, we can use the diagram to develop a branching, tree-like structure. After this has been done, we will see that there are 18 possible paths that one can follow through the tree (from left to right in the example we'll develop). Each path will be unique and will represent one of the possible scenarios specified by influence diagram. Since the outcomes specified at each node of the influence diagram were exhaustive, exactly one of these scenarios will (in our analytical judgment) happen.

To begin structuring a "tree" that represents our influence diagram, we first make use of the bottom nodes of the diagram. That is, we first use those nodes that have no (zero) incoming arrows. Since they have no incoming arrows, they (by the definition of the arrows) are not influenced by anything else and the relative likelihoods of the outcomes associated with them are ones we feel comfortable in assessing directly.

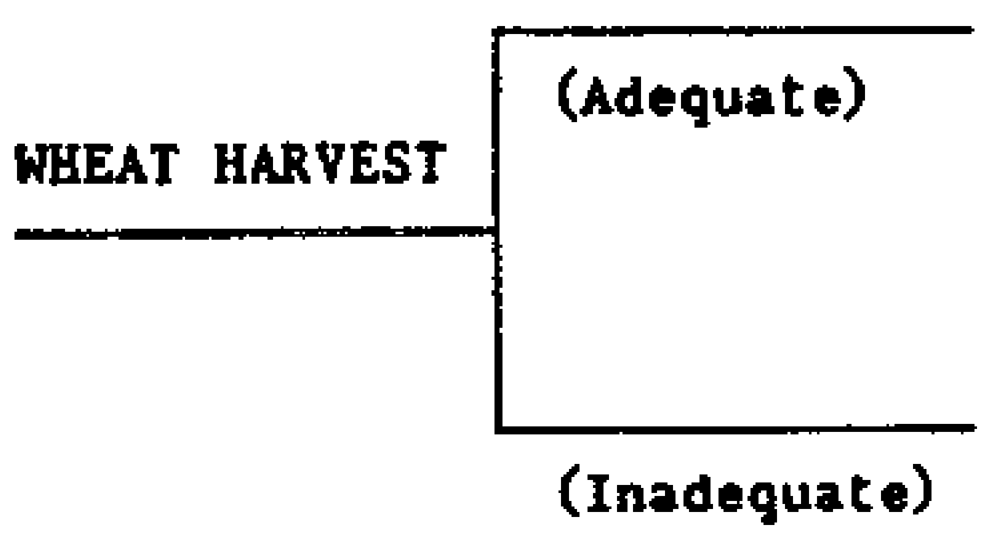
In the specific case of our example influence diagram, there is only one bottom level node (WHEAT HARVEST). Accordingly, we shall start with it. This we shall do (since we'll be working from left to right) by representing, at the left edge of the page, the WHEAT HARVEST event as an "event line" that branches, at its right-hand end, into two separate and distinct outcomes (Figure 29).

This completes work on the WHEAT HARVEST node. We next inspect the influence diagram once again to find the next node that can be added to the tree. Since we have already included the WHEAT HARVEST node in the tree diagram, we now can select nodes that have incoming arrows from the WHEAT HARVEST node only. We shall add the outcome "event line" and "outcome fork" for this newly selected node after <u>each</u> outcome branch from WHEAT HARVEST. This is because the outcomes associated with this new event will occur regardless of what size of wheat harvest is cbtained.

---

[32] The number of possible scenarios can be computed by multiplying together the number of possible outcomes for each event. In this case WHEAT HARVEST has 2, COST OF LIVING has 3, and ELECTION RESULT has 3. (2 X 3 X 3 = 18)

FIGURE 29 -- FIRST NODE OF TREE REPRESENTING INFLUENCE DIAGRAM



Using the above selection criteria, the COST OF LIVING node now qualifies. We will include it in two separate places in the tree, once after each outcome branch of the WHEAT HARVEST event. The influence diagram indicates, since COST OF LIVING has only one incoming influence arrow, that we feel comfortable evaluating the relative likelihoods of the COST OF LIVING outcomes <u>if</u> we know the outcome of the WHEAT HARVEST. The tree diagram gives us this information, since we know which WHEAT HARVEST outcome branch we are on and, hence, under which WHEAT HARVEST outcome assumption we are assessing COST OF LIVING outcomes.

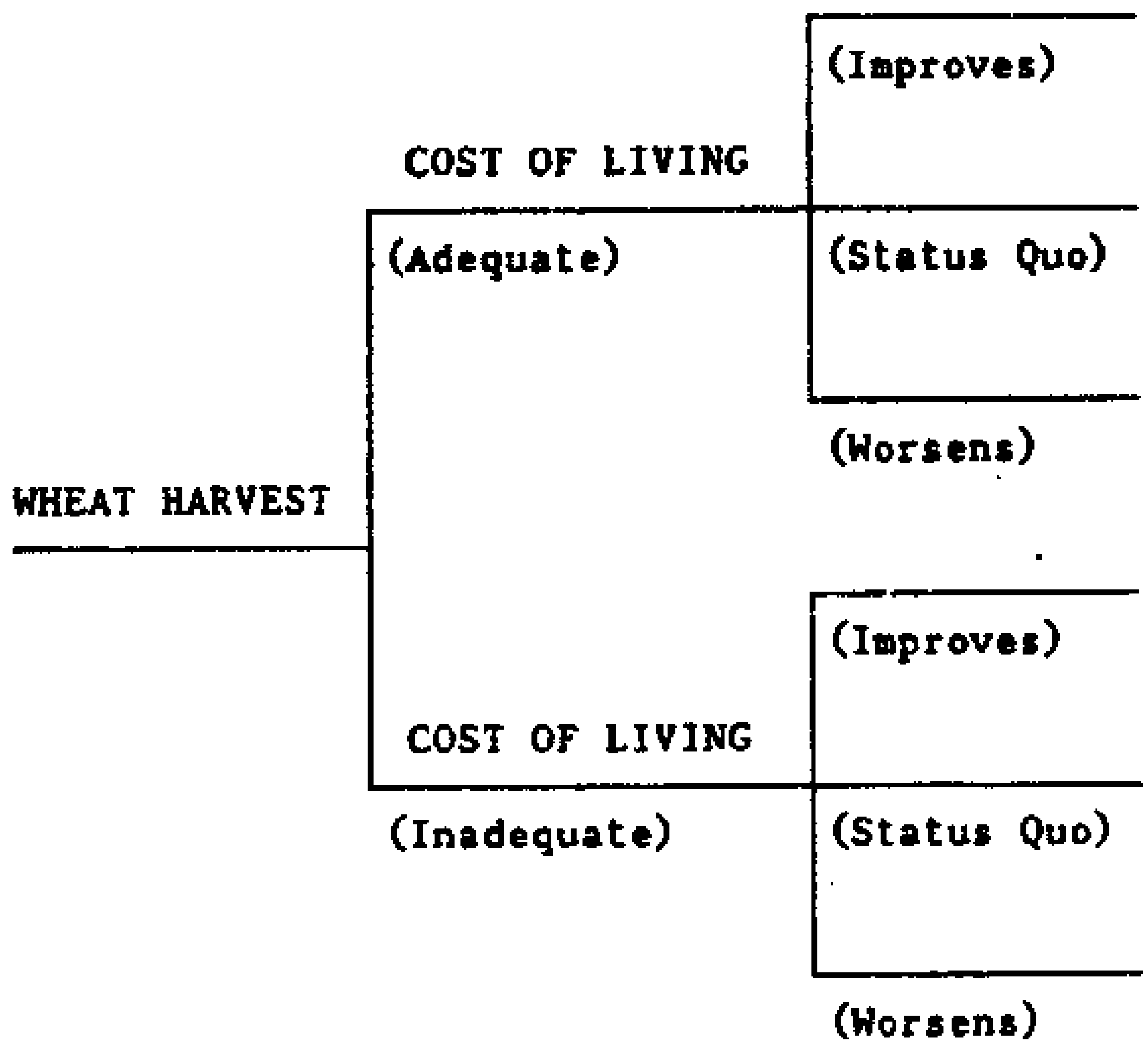
Said another way, when we assess the relative likelihoods of the COST OF LIVING outcomes, the tree will remind us of the conditions under which we are making the assessment. For, to get to the COST OF LIVING outcome forks from the start (left-hand side) of the tree, we must trace a path along either the Adequate or the Inadequate outcome branch of the WHEAT HARVEST event. To assess the relative likelihood of the COST OF LIVING outcomes, we simply <u>assume that the path we took is the actual state-of-the-world</u>.

The new tree (with the COST OF LIVING forks added) now looks that shown in Figure 30.

At this point let's introduce some minor bookkeeping shorthand to simplify the tree somewhat. Instead of labeling each branch with both "EVENT NAME" and "Outcome Name," we'll simply indicate the "EVENT NAME" above the tree with the understanding that all of the outcome forks located vertically below the given "EVENT NAME" are the forks describing that EVENT. This will revise the tree to that shown in Figure 31.

We now can return to the influence diagram to determine the next node to be entered into the tree. In the case of our example only one node (ELECTION RESULT) remains. Since we have already included in the tree the outcome forks for all of the other events, we can ignore the influence arrows from them to the ELECTION RESULTS node, and add this node to the

FIGURE 30 -- FIRST AND SECOND NODES OF INFLUENCE TREE



tree.[33] The result is Figure 32.

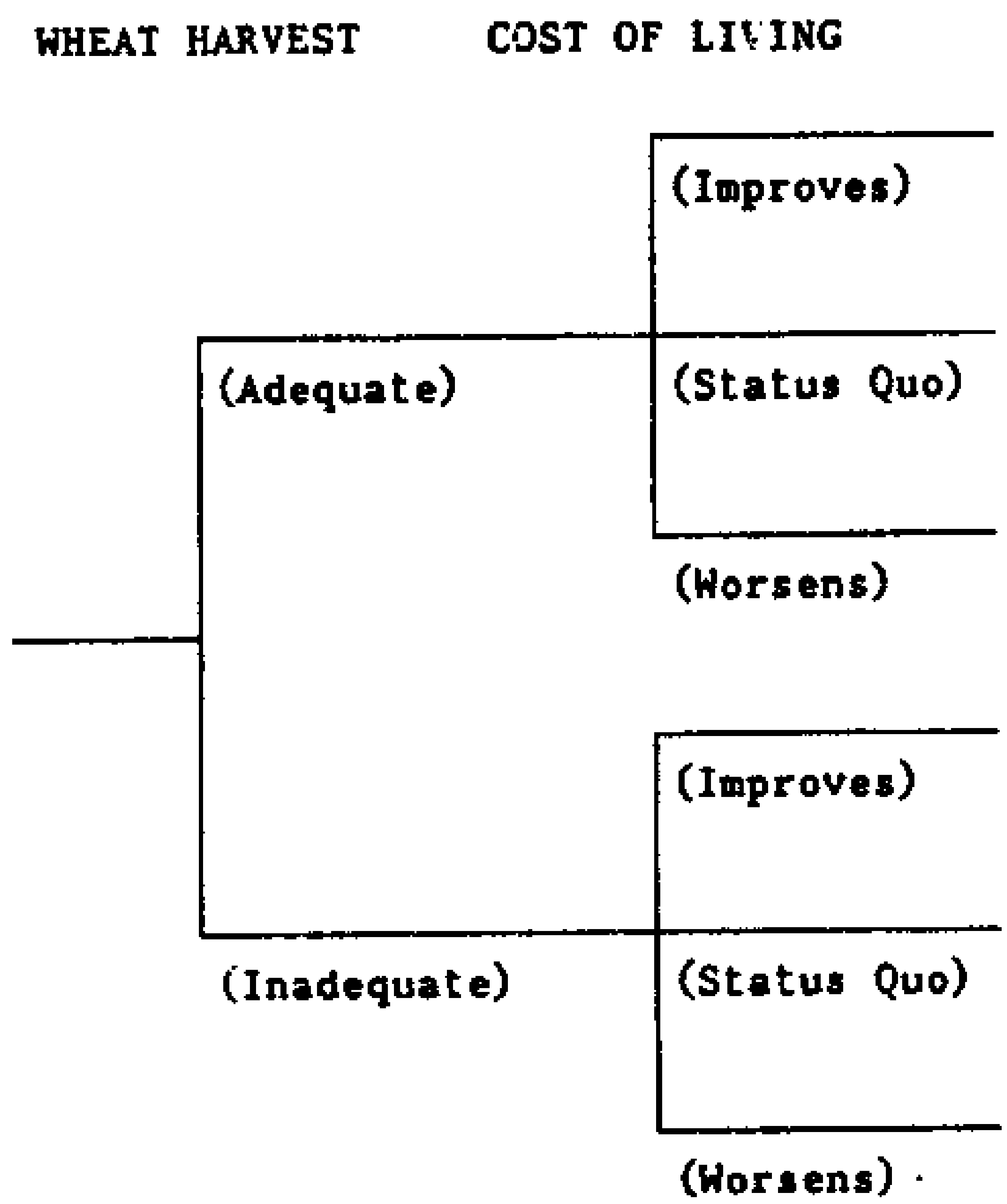
A list of the 18 possible scenarios is shown in Table 14. Each scenario is described on the tree by the unique path that can be traced to one of the 18 right-hand "end nodes." Note that there are 6 possible ways for each candidate to win.

Table 15 lists the 18 possible scenarios grouped together according to common election results. In assessing the probable election results, we need to determine which group of common-election-result scenarios is most likely. This can be done intuitively using the tree structure or the list of 18 scenarios as guides or checklists, or it can be done more explicitly using some simple graphical or numerical techniques.[34] These options,

----

[33] In the example used here, we had no choice as to which node to select next. Frequently this will not be the case. In those situations where multiple, acceptable bottom nodes exist, they can be incorporated into the tree in any order. This does not affect the final result in terms of either the number or specification of the possible outcome scenarios. You may find a given ordering of EVENT forks to be more or less convenient than another (and you should use the one you like best), but the technical accuracy of all tree variations is identical as long as the construction principles cited above are followed.

FIGURE 31 -- SIMPLIFIED INFLUENCE TREE



particularly the numerical one, will be discussed in some detail.

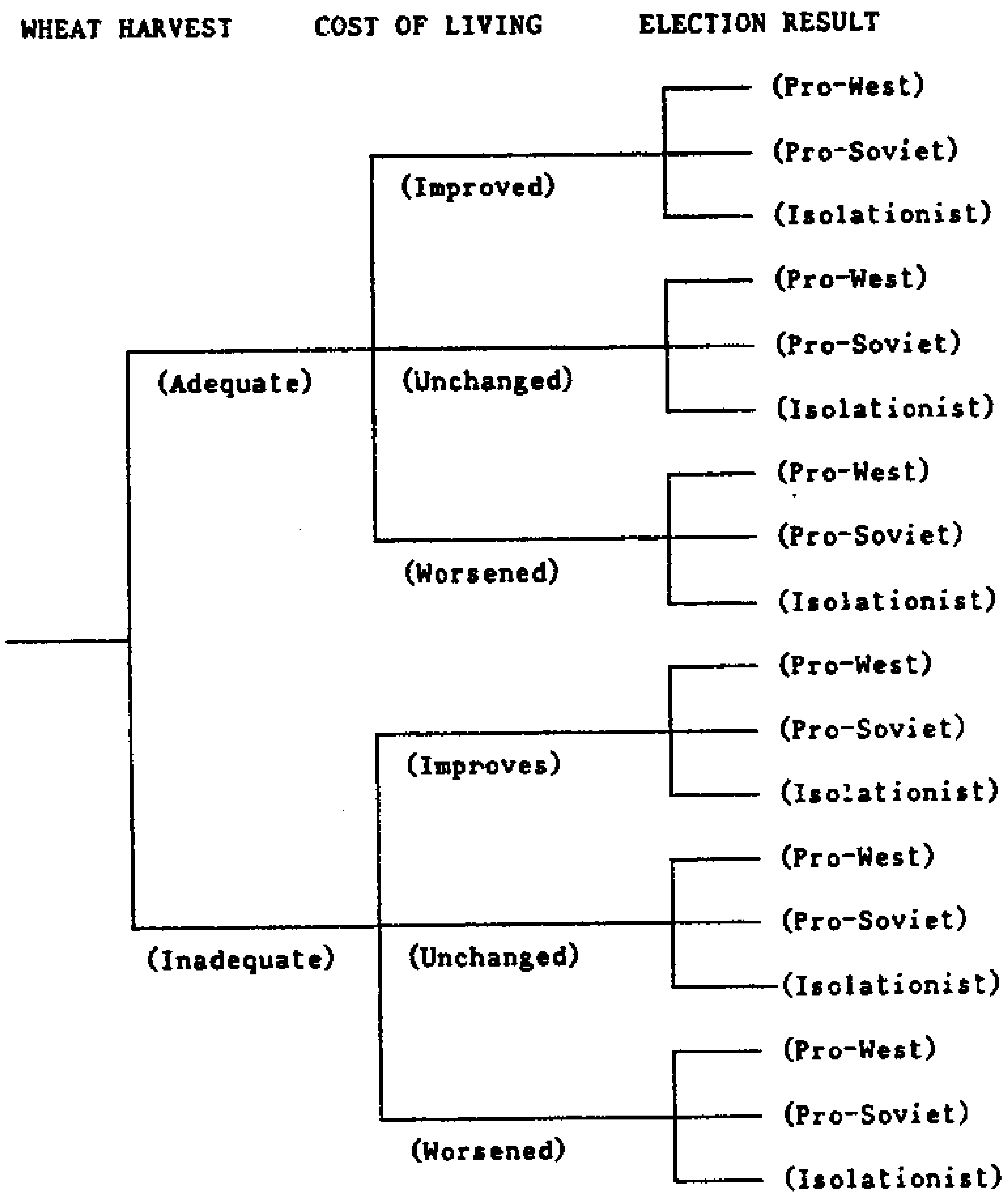
INTUITIVE ESTIMATES OF RELATIVE SCENARIO LIKELIHOODS

If one chooses to approach the task of assessing the relative likelihoods in an intuitive way, the following may be of some use:

---

[34] The graphical and numeric techniques are actually intuitive techniques. They rely on users to make subjective estimates of outcome likelihoods, as task which most people perform suboptimally. An extensive literature exists on subjective probability assessments. Interested readers could start with the Are You Sure pamphlet that can be obtained from the Information Science Center on (703)-351-3532, or with an article by Dick Heuer, Biases in Estimating Probabilities: The Availability Rule, Analytical Methods Review, March, 1980, pp.14-21.

FIGURE 32 -- COMPLETED SAMPLE INFLUENCE TREE



- Label each outcome branch on the tree with a descriptor (e.g. highly likely, improbable, etc.) that indicates your assessment of the chances of that specific branch happening relative to the chances of one of the other outcomes of the same node occurring instead. (Consider only your assessment at that node, not your assessment of reaching the node through the chain of outcomes defined by the outcome-path that leads to the node from the start of the tree.) Use this labeled tree as an aid to assessing the overall relative likelihoods of the common final outcomes (e.g.,

PROBLEM SOLVING TECHNIQUES

Probabilistic Reasoning

tree. The audit trail feature can aid future attempts to revise or update this assessment, since your current thinking has been documented fairly explicitly.

- As an alternative labeling scheme, label each branch with a likelihood descriptor (as above); but take into account the specific path that leads to the EVENT node that you are assessing. The labels you assign to each outcome at a given EVENT node should still reflect your assessment of their likelihoods relative to each other, but they now must reflect assumed "prior history" as well. For example, if you feel that an adequate harvest is unlikely, then none of the outcomes of any branch lying to the right of this "unlikely" outcome can be assigned any descriptors suggesting any likelihood higher in expectation. In fact, each outcome should be assigned descriptors implying less chance of happening than the "unlikely" Adequate Harvest outcome that is assumed to have preceeded it. (This reflects the fact that this procedure divides the "remaining" likelihood that reaches any given EVENT node among all of the outcomes that leave the node; so the total likelihood of all of a given node's outcomes cannot be greater than the likelihood of reaching that node via the outcome-path leading to it from the start of the tree.) The final 18 descriptors, therefore, indicate the likelihoods assigned to each of the 18 unique Outcome Scenarios that are specified by the tree. Thus, the "sum" of the final 18 descriptors should intuitively "equal certainty." That is, the final set of 18 descriptors must obey the same <u>mutually exclusive and exhaustive</u> requirements imposed earlier on the outcomes at each EVENT NODE. Therefore, we cannot have two (or more) of the final 18 scenario descriptors, each of which indicates a likelihood of occurrence that would make more than one of the scenarios the "ONE" expected to happen.

GRAPHICAL ESTIMATES OF RELATIVE SCENARIO LIKELIHOODS

Using the tree structure to assess graphically the composite, relative likelihoods of the three ELECTION RESULT outcomes can (compared to the intuitive techniques just discussed) help your overall assessment to be consistent with your feelings about the relative likelihoods of the outcomes at each of the EVENT forks in the tree. (Human beings tend to do poorly at composite assessments when approached in a totally unaided, subjective manner.)

The following are possible graphical approaches:

- Modify the tree structure developed above by weighting the thickness of the branches at each outcome node to represent your assessment of the relative likelihood of each of the outcomes. For example, if you feel that an adequate crop harvest is twice as likely as an inadequate one, make the Adequate branch twice as thick as the Inadequate branch. Do the same for each outcome node independently; except that remember that your assessment at each node should be made assuming that the true state-of-the-world under which you are making the assessment is that of the path of outcome branches leading from the start of the tree to that specific node that is being assessed currently. For example, when assessing the topmost ELECTION OUTCOME node, assume that there has in fact been an adequate wheat harvest and that the cost of living has improved. Use the line weights as a reminder or indicator of the relative likelihoods of each of the 18 different outcome scenarios. (This weighting scheme represents the graphical equivalent of the first intuitive, verbal-descriptor, labeling process suggested above.)

- To extend the graphical weighting scheme just discussed one step further, weight the lines in such a fashion that the total thickness of all lines leaving any given node equals the thickness of the line entering the node from the left (except, of course, for the very first node, where the initial line's width is up to you). The relative thicknesses of the lines leaving each node should reflect your assessment of the relative likelihoods of the respective outcome branches. If you follow this procedure, a useful thing will happen. Namely, the total thickness of all of the final outcome nodes will equal the thickness of the line leading into the starting node on the tree. By simply measuring and adding together the thicknesses of all of the final-node branches for each common (main-event) outcome, the relative likelihoods of the main-event outcomes can be determined directly. For instance, if you start the tree with a one-inch thick line, the sum of all eighteen final outcome branches will equal one inch also (to within measurement accuracy). If the summed thickness of the six Pro-West outcomes was one-half inch, the Pro-Soviet branches one-quarter inch, and the Isolationist branches one-quarter inch, then the Pro-West outcome would be about twice as likely as either of the other two outcomes (when considered individually), and about equally likely as either of the other two (the other two considered together). (This weighting scheme is the graphical equivalent of the second intuitive, verbal-descriptor, weighting scheme discussed above.)

NUMERICAL ESTIMATES OF RELATIVE SCENARIO LIKELIHOODS

The second graphical weighting scheme described in the final option above can be approached numerically. The numerical method has the advantages of being easier to change or update (which permits easier sensitivity testing -- a topic to be discussed below), being easier to use (particularly in the case of large tree structures), and being easily supported by computers (several software packages exist currently on most Community computer systems). But numerical methods also have the inherent disadvantage of implying (merely because NUMBERS are used) a false degree of accuracy. The numbers used below are precise (i.e., unambiguous in meaning), but they are no more accurate than the subjective feelings (usually expressed via the descriptor words used earlier) they are used to represent. This significant subtlety usually is lost on an audience (consumer) presented with numbers, and all too frequently is forgotten by the analyst employing numerical probability techniques.

To use numbers instead of weighted lines, one needs only to assign numbers between 0 (zero) and 1 (one) to each of the Outcome branches leaving each EVENT node in such a fashion that

- the sum of all of these "outcome weights" at each node add up to 1.0 (one), and
- the relative size of the numbers assigned to the set of Outcomes at each EVENT node reflect the relative likelihoods of occurrence as assessed by you.

If this is done, the set of numbers lying along each Outcome branch of the path describing one of the Outcome Scenarios can be multiplied together to determine that particular scenario's relative likelihood (as assessed by you) compared to to the likelihood of any other scenario path for which you perform the same calculation. Furthermore, you will find that if these relative likelihood numbers are calculated for all 18 final Outcome scenarios, they will sum to one (1.0). Therefore, the number calculated for each specific Outcome scenario represents directly the fractional chance (fraction of one) that the associated scenario has of occurring.

Since the sum of all of the Scenario likelihood numbers is one, the fractional number representing any given scenario can be multiplied by 100 to determine the percent chance that the scenario has of occurring. For instance, if the product of the Outcome weights leading to the "top-most" scenario had a value of 0.08, then that scenario would have been assessed a 8 percent chance of occurring. (Those familiar with probabilities will recognize that the 0.08 is a numerical expression of probability, since the sum of all possible outcome scenarios sum to one and probabilities are expressed as numbers between zero and one.)

The ability to assess probability (or percent chance) of occurrence directly from the number calculated for each outcome scenario works also for sets of final outcomes. That is, one can (as was suggested for the second graphical technique discussed earlier) sum the numbers (instead of the line thicknesses) for any set of outcome scenarios to determine the percent chance that any one of the outcome scenarios contained within that set will occur.

The tree shown in Figure 33 illustrates some of the things that have just been discussed. We again use the tree developed above (Figure 32). The only things that have been added are weighting numbers (having values between 0 and 1) to indicate a set of assessed relative likelihoods for the set of outcomes of each EVENT node. The numbers at the right-hand side of the tree indicate the relative likelihood of occurrence of each of the 18 Outcome Scenarios (as calculated by multiplying together all of the weighting numbers lying on the outcome branches defining the scenario they label).

FIGURE 33 -- TREE USED TO ASSESS OUTCOME LIKELIHOODS NUMERICALLY

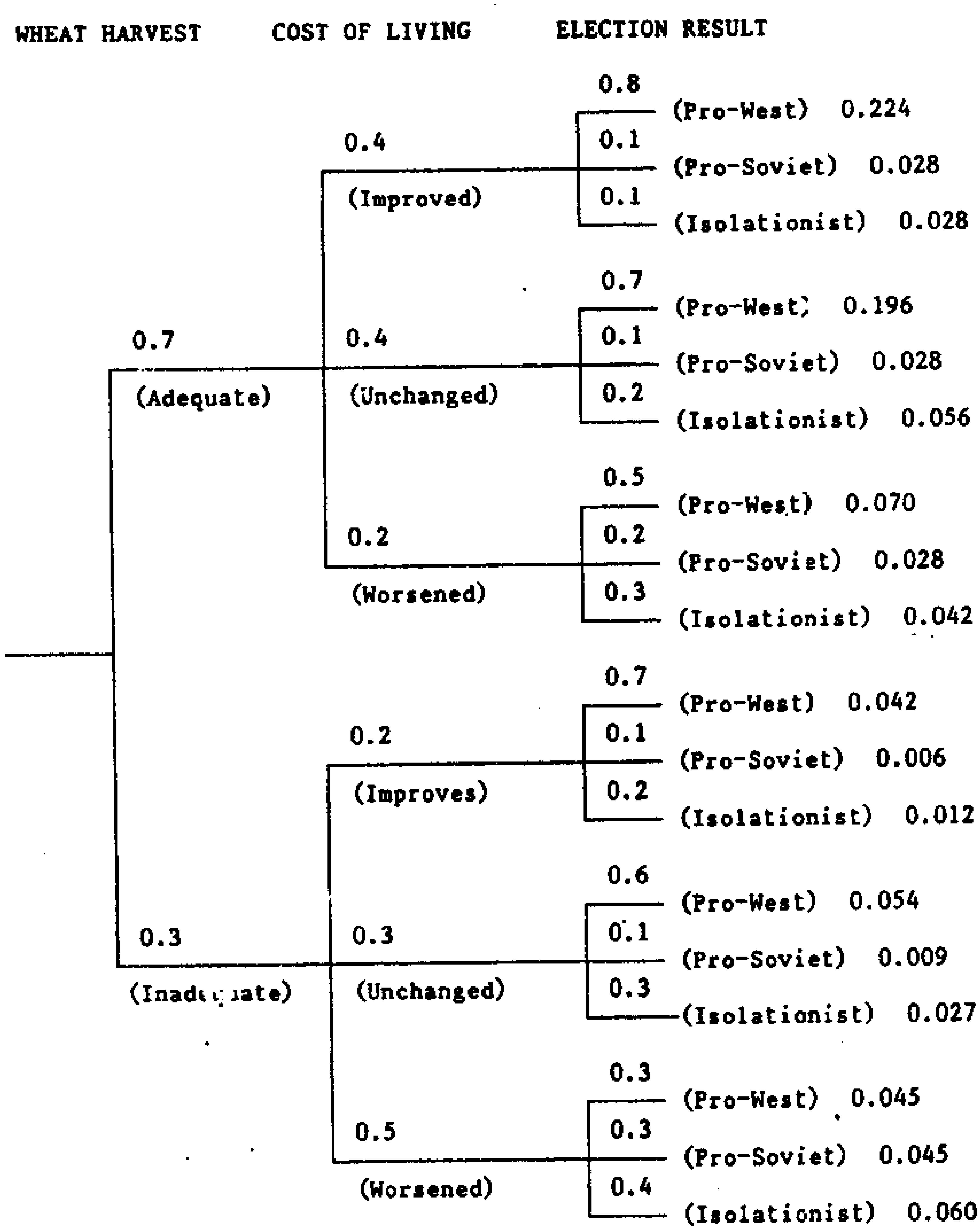


TABLE 16 -- SUMMED LIKELIHOODS OF COMMON OUTCOMES

Pro-West	Pro-Soviet	Isolationist
0.224	0.028	0.028
0.196	0.028	0.056
0.070	0.028	0.042
0.042	0.006	0.012
0.054	0.009	0.027
0.045	0.045	0.060
0.631 = 63 %	0.144 = 14 %	0.225 = 23 %

Table 16 shows the summed likelihood of occurrence of each of the three mutually exclusive Election Outcomes (Pro-West, Pro-Soviet, and Isolationist). Based on the assessments used for the individual Event node outcomes comprising the tree, the summations in the table tell us that a Pro-West election result is about four times as likely to occur as a Pro-Soviet result, and about three times as likely as an Isolationist result. We also could use the summed figures to note that a Pro-Soviet result only has about one chance in seven (14 out of 100) of occurring. The Pro-West result is about twice as likely to happen (based on our assessments) as the composite chance of anything else happening instead.[33]

These computations demonstrate how the numerical tree method can provide prospective on the question of primary interest (the election outcome) while permiting the concentration of evaluation efforts on the events that we feel will determine, or at least influence, the event that is of foremost interest. But, we can even do better; for we can use the same procedure to check the importance of uncertainties we had in making the individual assessments for the numbers entered into the tree. We thereby can assess the usefulness of spending analytical time studying any given issue more fully. (This method is discussed next under the topic heading of Sensitivity Analysis.) Similarly, this same procedure can be used to assess the impact that a new piece of information should have on our previous judgment of relative scenario or outcome likelihoods. (See the section below on Updating Probability Estimates.)

---

[33] It is worth noting again at this point the same words of caution with which this section on numerical estimation was begun. Namely, that although the final relative likelihood numbers we have calculated are quite precise in meaning, they are in fact no more accurate than the words (qualifiers such as "highly likely" or "improbable") that they replaced. The use of numbers simply has helped us relate several separate likelihood estimates in a consistent and documented manner.

SENSITIVITY ANALYSIS

Let's examine the first of these additional capabilities (evaluating the relative importance or impact of changes in the values we assign to influencing event outcomes). Assume, for instance, that we were uncertain about the relative likelihoods we assigned above to the COST OF LIVING Outcomes. We used our best estimates, but feel that the more optimistic and more pessimistic sets of numbers shown in Table 17 might be reasonable also.

TABLE 17 -- LIST OF PESSIMISTIC AND OPTIMISTIC EXTREMES

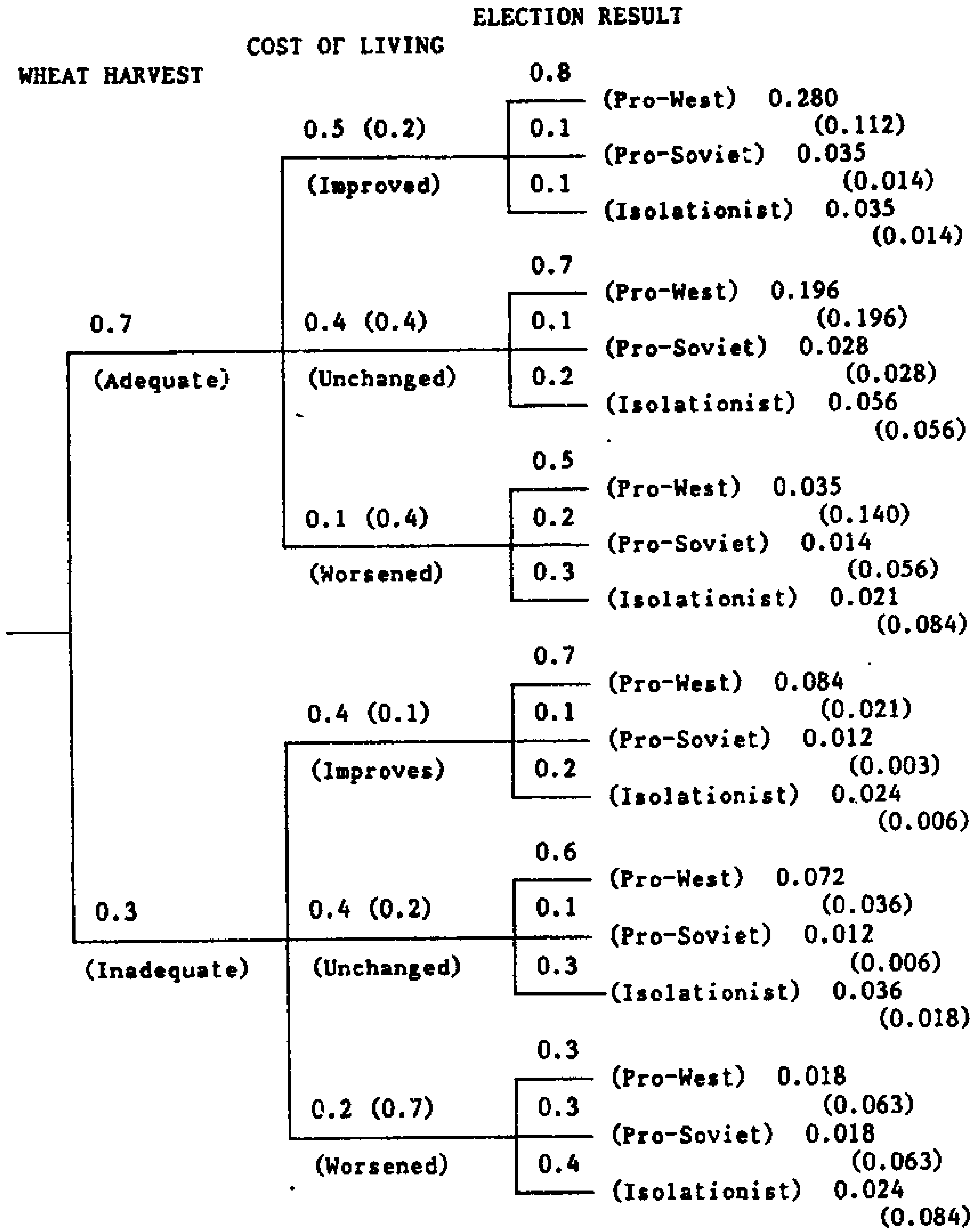
	COST OF LIVING		
	Optimistic		Pessimistic
Adequate Harvest	0.5	Improves	0.2
	0.4	Status Quo	0.4
	0.1	Worsens	0.4
Inadequate Harvest	0.4	Improves	0.1
	0.4	Status Quo	0.2
	0.2	Worsens	0.7

To assess the extent to which changing these numbers would affect our overall assessment of the Election Outcomes, we simply replace the COST OF LIVING numbers we had in the tree with the two sets proposed, and then calculate the net impact of each new set. Figure 34 shows the tree with the optimistic and (in parentheses) pessimistic numbers. The scenario likelihoods listed at the right of the tree show the pessimistic outcomes enclosed in parentheses also. Table 18 shows the summarized results for the three ELECTION RESULT outcomes for the optimistic, best-guess, and pessimistic assessment numbers.

As can be seen from Table 18, the optimistic and pessimistic calculations indicate that the likelihood of a PRO-WEST Outcome ranges from about 69% (for an optimistic appraisal) to about 57% (at the pessimistic extreme). Our best guess happens to fall at the middle of this range; and we can therefore say that our assessment of a PRO-WEST Outcome (as we have structured it here) is about 63% (plus or minus about 6%) likely.

The variation of about 6% either side of our initial guess may seem big or small to you. Its size may or may not have surprised you compared to what you expected the results of these calculations to be. The point is that by performing the calculations, we were able to evaluate the sensitivity of our Outcome likelihood estimates to changes in our assessment(s)

FIGURE 34 -- SENSITIVITY ANALYSIS USING A TREE STRUCTURE



of an influencing event(s). We also have an audit trial of what our assessments of each influencing outcome scenario were, and a well-defined structure of how we viewed the problem. Either the structure or the assessment numbers can be changed (revised or updated) any time we choose.

TABLE 18 -- SUMMARIZED PESSIMISTIC-OPTIMISTIC OUTCOME RANGE

	PRO-WEST	PRO-SOVIET	ISOLATIONIST
OPTIMISTIC	68.5%	11.9%	19.6%
BEST GUESS	63.0%	14.0%	23.0%
PESSIMISTIC	56.8%	17.0%	26.2%

Their explicit format makes this process relatively straight forward. It also makes it convenient to share our analysis with coworkers or consumers for either explanatory or advice-seeking purposes.

UPDATING PROBABILITY ESTIMATES

The updating uses of probability trees like the one we constructed go beyond simply revising numbers that have been assigned to the tree in the largely intuitive manner we used above to perform a sensitivity analysis. A procedure known a Bayesian updating can help us to revise our assessments in a way that is itself well-structured and explicit. It deals with using new evidence about events that relate to the EVENTS detailed in our tree, but which are themselves not included in the tree structure.

Bayesian calculations can be used in several different forms. We'll use the Likelihood Ratio approach.[34] But regardless of the form used, Bayesian techniques require one to "think backwards" about the problem from the way most people approach problem solving.

For example, assume that we get some new information (usually referred to as an Indicator or an Indicator Event) that should help us decide among or re-evaluate our likelihood assessments of one of the Event Nodes in a probability tree. Let's say we just learned (a few day's after having completed work with the tree developed above) that the Agricultural Minister of our fictitious country has contracted with the Canadians to buy 1 million metric tons (10% of their projected yearly domestic need) more wheat this year than we would have expected based on previous import practices. Typically, when given such information, a person will try to incorporate this new information in his or her intuitive data bank and try to reassess the likelihoods previously put on the Wheat Harvest outcomes of the tree (Figure 33) developed above. What Bayesian procedures require is for one instead to assess the likelihood of receiving the news about the

---

[34] Both the likelihood ratio and a second method that calculates probabilities directly are discussed (using mathematical probability notation) starting on page 177 of this Handbook.

Indicator if one assumes, one at a time, that each Outcome of the Event node that is being reassessed is in fact the actual State-of-the-World. For the case of our grain harvest, we first assess the likelihood of receiving the reports of the Agricultural Minister's large grain order assuming that the actual Wheat Harvest will be adequate. (The number we assign here should be scaled between zero and one -- or 0 and 100% -- where 0 means no chance of occurrence, and 1.0 -- or 100% -- means absolute certainty of occurrence. Let's assume we assigned this a 0.6 on the basis that the market was favorable to purchasers and the Minister himself probably does not know for certain that the still pending wheat harvest will be large.) We next must assign a likelihood to the report about the large wheat order assuming that the harvest will be inadequate. (Let's assume that we assign a value of 0.9 based on the assumption that the Minister has good indication that the grain will be needed to replace a expected, inadequate domestic harvest.)

What one does using Bayesian procedures is to take the ratio of these two assigned numbers (the procedure also works if more than two Outcomes are being reassessed) and multiplies it times the ratio of the prior likelihood values that were listed in the tree. (If the Event node we were revising had appeared in several different places in the tree, we would do this multiplication separately for each set of likelihood values assigned to that Event node's occurrence in the tree.) In our case, the Wheat Harvest event node occurs only one time, and has prior probabilities of 0.7 and 0.3 assigned to the Adequate and Inadequate branches, respectively.

At this point we multiply the 0.7 and the 0.3 by the 0.6 and the 0.9, respectively.

$$\frac{0.7}{0.3} \times \frac{0.6}{0.9} = \frac{0.42}{0.27}$$

This gives a new set of odds for our revised estimates of the outcome likelihoods.[37] The odds can be changed to the likelihood values (probabilities) we need for use in the tree quite easily. We simply divide each of the members of the revised odds by the sum of all of the members (in mathematical terms, we normalize them). For our example, the new likelihood value for an Adequate Harvest (called the posterior probability in Bayesian terminology) is:

$$\frac{0.42}{0.42 + 0.27} = \frac{0.42}{0.69} = 0.61$$

and the revised likelihood value for an Inadequate Harvest is:

---

[37] Remember that the likelihood values assigned to each outcome set at every node node sum to 1.0. This normally will not be the case for the revised odds produced by this Bayesian technique.

$$\frac{0.27}{0.42 + 0.27} = \frac{0.27}{0.69} = 0.39$$

At this point we can revise the 0.7 and 0.3 relative likelihood estimates used in Figure 33 (for the Adequate and Inadequate Harvest outcome branches, respectively) to 0.61 and 0.39 respectively; and then recalculate the relative likelihoods of the 18 Outcome Scenarios. The new Outcome Scenario likelihood numbers will permit us to determine the extent to which our overall assessment of the ELECTION RESULT Outcomes has been affected.

Note that the new likelihood values (0.61 and 0.39) maintain the same ratio the revised odds had, but that they also sum to 1.0. If at this point in time another indicator event occurred, we would simply repeat the same procedure. This time around, we would use the 0.61 and the 0.39 likelihood values that we just calculated as our current (or prior) odds. If several Indicator events are reported together, the order in which we consider them for purposes of updating the likelihood values in the tree will make no difference on the final outcome likelihoods we compute. (This is true because multiplication is commutative.) In fact, to save time one can assess the likelihood ratios involved for each of the indicators and perform just one long multiplication (normalizing the single set of odds that result) instead of performing each multiplication operation separately, and normalizing the resulting odds at each step.

Note also that the extent to which the initial likelihood estimates are changed by multiplying them by the Bayesian likelihood ratio depends directly on the difference(s) in magnitude of the probabilities assigned to the likelihood ratio. That is, if one estimated a likelihood ratio of:

0.7 : 0.3 : 0.2

for some indicator event influencing our judgements about the three election outcomes we considered earlier, then we can tell immediately that using this likelihood ratio to update our estimates of the election outcomes will decidedly increase our estimated chances of the first of the three outcomes occurring relative to the other two, since the 0.7 corresponding to the first outcome is large (over twice the magnitude) of either of the other two outcomes. This information tells us that an indicator that produced a likelihood ratio such as the one suggested would be a good discriminator among possible event outcomes. An indicator that produced a "flat" likelihood ratio (e.g., 0.5 : 0.45 : 0.5) would provide little discrimination among the potential outcomes. Since one can assess the likelihood ratios for potential indicators, this gives us a tool to assess the potential usefulness of spending time or other resources to to obtain information about some indicator event against which one is considering writing requirements.

ANALYSIS ACTIVITIES

Probabilistic Reasoning

PROBABILITY THEORY

TERMINOLOGY

Probability

A probability defines the chance that a given event will occur, expressed as a fraction from zero (signifying an impossible event) to one (signifying an event that will definitely occur), or the corresponding percentage from 0 to 100. The symbol $P(A)$ will be used to mean the probability of event A.

Independent Events

Independent events are events which do not influence the probability of occurrence of each other. For example, the probability of obtaining a head on the second toss of a coin is independent of the result of the first toss.

Mutually Exclusive Events

Mutually exclusive events--events which cannot occur simultaneously, e.g., heads and tails are mutually exclusive on a single flip of a coin.

Conditional Probability

A conditional probability gives the probability of occurrence of one event, given that some other event has occurred or will occur. $P(A|B)$ is read "the probability of A, given that B has occurred or will occur." For example, we may assess the probability that the Redskins will win on a given weekend (event A) as 0.80 or $P(A) = 0.8$. But we may think that the probability of them winning if Joe Theismann is unable to play (his injury being event B) is lower, say 0.60 or $P(A|B) = 0.6$. If on the other hand, the events are independent, then they do not influence the probability of each other. In that case, the conditional probability of one, given the other, is still equal to the original probability of the first. For example, the probability that the Redskins will win (0.80) is independent of whether or not I go to the game (my presence being event C). So, $P(A|C) = P(A) = 0.8$.

Joint Probability

A joint probability is the probability that both events A and B will occur. If A and B are independent, it is simply the product of their separate probabilities. For example, the probability of a flipped coin landing heads is 0.5; the joint probability of two heads in a row is $0.5 \times 0.5 = 0.25$

PROBLEM SOLVING TECHNIQUES

Probabilistic Reasoning

Odds

Odds indicate the chance that a given event will occur, expressed as a ratio rather than a percentage or fraction as in probabilities. If the probability of an event occurring is 55%, for example, the odds for its occurrence are 55 to 45 (100 minus 55 = 45) and the odds against its occurrence are 45 to 55, or 11 to 9 for, and 9 to 11 against, respectively.

Factorial (!)

Factorials are the product of successive integers, always ending with one. For example, 7! (read "seven factorial") equals:

$$7 * 6 * 5 * 4 * 3 * 2 * 1 = 5040$$

By definition, 0! = 1.

> Permutations and combinations are used for certain types of exact theoretical probability calculations.

Permutations

Permutations are the arrangements that can be made from "n" distinquishable objects, taken "r" at a time, when the order of the arrangement is important. For instance, all of the possible rankings of the top 5 students in a class of 20 would be the permutations for 20 objects taken 5 at a time. The formula for calculating the number of possible permutations is:

$$P = \frac{n!}{(n - r)!}$$

Combinations

Combinations refer to the sets of objects that can be formed from "n" objects taken "r" at a time, when order is not important. For instance, the number of groups of students that could comprise the top 5 students out of a class of 20 students, if one does not care what the individual rankings of the 5 students are, would be the set of combinations of 5 students that could be formed from a class of 20 students. The formula for determining the number of possible combinations is:

$$C = \frac{n!}{r!\,(n - r)!}$$

ANALYSIS ACTIVITIES

RULES OF PROBABILITY

Range of Probabilities

$$0.0 \leq P(A) \leq 1.0$$

Summation Rule

The probability of event A or event B occurring is equal to the probability of A occurring plus the probability of B occurring, i.e., P(A or B) = P(A) + P(B), provided A and B are mutually exclusive events. If A and B are not mutually exclusive then:

$$P(A \text{ or } B) = P(A) + P(B) - P(A \text{ and } B)$$

where P(A and B) means the probability of A and B occurring simultaneously.

Multiplication Rule:

The multiplication rule is applied to joint probabilities. The probability that event A and event B will both occur is the product of their individual probabilities. If the events are independent, then their joint probability is:

$$P(A \text{ and } B) = P(A) * P(B).$$

If, however, they are not independent events, then their joint probability is:

$$P(A \text{ and } B) = P(A) * P(B|A)$$

if P(B) depends upon the occurrence of event A or P(A and B) = P(B) * P(A|B) if P(A) depends upon the occurrence of event B.

Rule of Complements

$$P(A) + P(\bar{A}) = 1$$

where $\bar{A}$ means "not A."

PROBLEM SOLVING TECHNIQUES

Probabilistic Reasoning

BAYESIAN ANALYSIS

Bayesian Analysis is a formal method for using incoming data to modify previously estimated probabilities. Each new piece of information may be evaluated and combined with prior historical or subjective assessments of the probability of an event in order to determine whether its occurrence has now been made more or less likely, and by how much. Bayesian analysis can also be used to compute the likelihood that the observed data are attributable to particular causes. One advantage claimed for Bayesian analysis is its ability to blend the subjective probability judgments of experts with historical frequencies and the latest sample evidence.

ODDS-LIKELIHOOD RATIO METHOD

Bayesian updating of probability estimates can be done in several ways. One of the easiest is the odds-likelihood ratio method. To use it, one expresses the current (usually referred to as prior) probabilities of the unknown state-of-nature that is of interest in their odds form. For instance, if there are 3 possible states-of-nature -- P(A), P(B), and P(C) -- having prior assessed likelihoods of 0.3, 0.6 and 0.1 respectively, then we would express them in ratio form as odds:

$$0.3 : 0.6 : 0.1$$

or

$$3:6:1$$

If some event occurs, (or does not occur) the likelihood of which we feel was dependent on the state-of-nature of our main event of interest, then we can use this new "indicator" event to update our set of prior odds. This is done by assessing the likelihood of the indicator event happening, in the fashion that it did happen, assuming in turn that each of the unknown states-of-nature is true. That is, we need to assess the conditional probabilities of the indicator event for each of the unknown states-of-nature of the event of main interest.

The conditional probabilities are then expressed in ratio form also, and the two sets of probabilities (the "prior" and the "conditional" probabilities) are multiplied together (pairing each state-of-nature with the corresponding conditional probability for the indicator event). If "I" is the indicator event, then:

$$P(A) * P(I|A) = P'(A)$$
$$P(B) * P(I|B) = P'(B)$$
$$P(C) * P(I|C) = P'(C)$$

gives us a new (posterior) set of odds for the state-of-nature probabilities:

$$P'(A):P'(B):P'(C)$$

These new odds (the posterior odds) can then be normalized (by dividing each P'(i) by the sum of all of the posterior odds) to obtain our revised posterior probabilities.

For example, if we had the following prior and conditional probabilities:

$$P(A) = 0.3 \qquad P(I|A) = 0.4$$
$$P(B) = 0.6 \qquad P(I|B) = 0.2$$
$$P(C) = 0.1 \qquad P(I|C) = 0.8$$

then:

$$P(A) * P(I|A) = 0.3 * 0.4 = 0.12$$
$$P(B) * P(I|B) = 0.6 * 0.2 = 0.12$$
$$P(C) * P(I|C) = 0.1 * 0.8 = 0.08$$

giving us posterior odds of:

$$0.12 : 0.12 : 0.08 = 3:3:2$$

Our posterior probabilities then become:

$$P(A) = P(B) = 3/(3+3+2) = 3/8 = 0.375$$

$$P(C) = 2/8 = 0.25$$

PROBABILITY METHOD

If one is comfortable in assessing directly both the conditional likelihood of an indicator given a particular state-of-nature, and also of assessing the overall likelihood of the indicator for all possible states-of-nature, then the following formula can be used to calculate a revised likelihood for a single state-of-nature (hypothesis).

$$P(H|I) = \frac{P(I|H)\ P(H)}{P(I)}$$

Where:

P(H|I) = the posterior (revised) probability of the hypothesis.

P(I|H) = the conditional probability of the indicator assuming the hypothesis is true.

P(H) = the prior probability of the hypothesis.

P(I) = the overall (non-conditional) probability of the indicator.

For example, if some hypothesis "A" had as an initial (prior) probability:

$$P(A) = 0.4$$

and some indicator event occurred for which we felt that there was only a 20% chance of the indicator happening if hypothesis A were true:

$$P(I|A) = 0.2$$

but that if we made no assumptions about the actual state-of-nature, then there was a 80% chance of the indicator occurring:

$$P(I) = 0.8$$

Then, we can calculate a revised (updated) probability for hypothesis A as follows:

$$P(H|I) = \frac{P(I|A)\ P(A)}{P(I)} = \frac{(0.2)\ (0.4)}{(0.8)} = \frac{0.08}{0.8} = 0.1$$

SUBJECTIVE ESTIMATES OF PROBABILITIES

This section has been adapted from a portion of the HANDBOOK OF FORECASTING TECHNIQUES, Part II, Description of 31 Techniques, prepared for the U.S. Army Engineer Institute for Water Resources, Fort Belvoir, VA. The work was done under Contract No. DACW31-75-C-0077, August 1977.

ABSTRACT

The subjective component is incorporated in every estimate of future probability, and consists of that basis for the weighting of respective outcomes to which no numerical basis can be assigned. As a method, however, the term refers especially to those procedures in which the subjective component is alleged or acknowledged to be of major or crucial importance to the forecast contents. Subjective estimates of probabilities are routinely relied upon in trivial, moment decision-making and in the vast, ultimate questions for which no objective appraisal is yet feasible. The subjective approach in fact has dominated most decision-making until the recent past, and undoubtedly dominates decision-making for most people in the world even at present. Many forecasting methods of recent vintage basically are systematic attempts to make subjective estimates of probabilities more explicit so that the implications of the multiple estimates can be determined consistently via objective mathematical techniques. This does not, of course, make the implications themselves objective, it simply makes them consistent with the initial subjective estimates.

DEFINITION

No potential future outcome is absolutely certain until it has been actualized in the present. In that strict sense, all estimates of probability are subjective. The concept of probability, however, implies that potential outcomes can be estimated and expressed in quantitative or at least metric terms--"the odds are 60 to 40 against," for example. The subjective component of any probability estimate consists of that weighting assigned to probabilities whose weight itself cannot be explicitly counted or measured--"my experience and intuition tell me that this outcome is less likely than that," for example. Every estimate of probability includes some component of subjectivity. The method specifically referred to here, however, includes all cases where the subjective component is alleged or acknowledged to be major and crucial in determining the content of a forecast.

This approach assumes that past experience can be and often is linked rationally and properly to the forecaster's ability to predict future developments and events. It is further assumed that some experience--in many cases, much of the most valuable experience--has not been and perhaps even never can be applied or described in numerical terms. For example, on what numerical basis might an art or music critic predict whether or not a current work of art will be recognized as such by future generations? The ultimate entrepreneurial risks in every enterprise are similar to judgments about works of art in this respect. In such situations, subjective, intuitive experience is invoked in making forecasts, either as a supplement or as the only credible basis available.

MAIN USES

Resort to subjective estimates of probabilities of future outcomes is determined primarily by the scale of cost, risk, and complexity involved in a given situation. In the making of routine, trivial, minute-by-minute decisions often it is neither necessary or worthwhile to strive for objectivity. In ordering lunch, selecting a suit of clothes, or choosing a route through the park, one's intuitive preferences are usually adequate. In speculating about the prospects for peace, humanity, religion, or the state objective estimates often are clearly inadequate and not worth making. It is only at intermediate levels of cost, risk, and complexity that it may seem worthwhile to invest in studied, objective estimates. State and national budgets, corporate diversification plans, manpower and training programs--these are examples of intermediate activities in which objective estimates of probability are attainable and useful. In a special sense, then, it is fair to say that we use subjective estimates of probabilities to settle both the smallest and the largest issues and uncertainties, reserving objective estimates for intermediate matters.

LIMITS AND CAUTIONS

Subjective estimates of probability are adequate so long as they are acceptable to all parties concerned or so long as objective events do not create unmanageable difficulties. There is a Great Pumpkin so long as Charlie Brown believes in Him and so long as He actually does appear on Halloween in a form even the skeptics must acknowledge. When these conditions are not met--as perhaps at present when people are ceasing to believe in the possibility of full employment because of current unemployment rates--grave, insurmountable crises may rapidly develop.

OTHER TECHNIQUES

As suggested above, subjective estimates of probabilities are used when more objective estimates are unnecessary or inaccessible. It should be noted that no estimation procedure is either purely subjective or purely objective. Any forecast topic has many aspects, and the question always must be how many and which aspects can be treated objectively or subjectively.

PROCEDURES

Unsurprisingly, considering how ancient and universal subjective estimation is, procedures for making subjective estimates are subjects of constant development, experiment, and testing. Subjective estimates are adequate only to the extent that they can be shared and are convincing; these conditions tend to be met only to the extent that estimation procedures tend to be explicit and logical and plausible, if not objective. It is in pursuit of these objectives that constant experimentation proceeds with the regularization of subjective procedures for estimating probabilities.

Given the range and diversity of subjective estimation methods, the corresponding range and diversity of products is commensurately large and diverse. The range varies between simple assertion--"I believe (for whatever reasons, unstated) that such-and-such will happen"--to elaborate, computerize models in which subjective weights have been quantized and cross-correlated in complex ways at many levels.

LEVEL OF CONFIDENCE OF RESULTS

Those who sponsor and those who offer these kinds of forecasts profess great confidence in them, because the forecasts are used to justify decisions and actions, past or pending. Beyond these parties, also, the level of confidence acknowledged in a subjective forecast usually depends on the values and views of those concerned. Where the forecast serves or can be made to serve positions established on other grounds the forecast is accepted with great confidence. Where the forecast attacks previously established positions it is ignored, attacked, or minimized.

PROBLEM SOLVING TECHNIQUES

Probabilistic Reasoning

COMMUNICABILITY OF RESULTS

Communicability is an interesting facet of subjective forecasts. A subjective forecast cannot serve its intended end unless it persuades to some extent those parties which are on record as hostile or skeptical. Well documented subjective forecasts typically permit major premises, major conclusions, and the chain of reasoning which links the two to be presented clearly.

SPAN OF FORECAST

Of all the aspects of forecasting by method, perhaps the most subjective of all is the estimate of how far into the future it is feasible to conjecture with credibility. Such being the case, subjective estimate forecast time horizons vary literally between the next sixty seconds and the remainder of the history of the universe, dependent purely on the topic, the forecaster's opinion, and the credibility of his audience. Given topics tend to have a characteristic time horizon, however: economic forecasts tend to deal with the next three to eighteen months, technological forecasts tend to deal with the next three to thirty years, etc.

RESOURCES NEEDED

Depending on the application, the topic, and the audience the range of resources literally varies between the minimum of one person with a strong opinion and a maximum of many years of opinion-poll data from normalized samples.

ASSESSING CONSTRAINED SYSTEMS

LINEAR PROGRAMMING

TERMINOLOGY

Objective Function
A mathematical statement of whatever it is that one desires to maximize (e.g., profit) or minimize (e.g., cost). It is expressed in terms of the unit value (cost, profit, etc.) of the system variables affecting the thing to be maximized or minimized.

Constraint
An equation or inequality that describes the feasible limits of values that the system variables can assume.

Feasible Solution
Any set of values for the system variables that do not violate the limits expressed by the constraint equations.

Optimal Solution
That solution (or, in rare instances, set of solutions) that optimize (maximize or minimize) the value of the objective function.

Nonnegativity constraints
Constraints on the variables that do not permit them to go negative (e.g., one cannot use a negative number of trucks to minimize shipping costs). Nonnegativity constraints normally are default settings (are automatically imposed) by most computer software.

Linear Program
A mathematical model consisting of a linear objective function, linear constraint equations, and (normally) nonnegative variables.

Linear Equations
Mathematical expressions in which the variables are contained in separate terms and are raised to the first power only.

Feasible Region
The set of all feasible solutions.

Slack variable
A dummy variable added to the right side of a less-than-or-equal-to constraint to convert it into an equality. Slack variable can usually be interpreted as the amount of unused resource.

PROBLEM SOLVING TECHNIQUES
Assessing Constrained Systems

Surplus Variable
A variable subtracted from the left-hand side of a greater-than-or-equal-to constraint to convert it into an equality. Surplus variables usually can be interpreted as the amount per variable unit that it would cost to deviate from the optimal solution.

TWO VARIABLE PROBLEMS

Linear programs are simply systems of linear equations or inequalities that are solved in a manner that yields as its solution an optimum value. The optimum value is determined based on some goal statement (provided to the program in the form of a linear objective function). Algebraic methods for obtaining the optimum solution are the same for two-variable and multi-variable problems, but our first example will be a two variable problem since graphical solutions (which can be visualized) can be obtained also. This should serve to help you understand conceptually why and how optimal solutions can be obtained.

To obtain a graphical solution one plots all of the constraint equations on a two dimensional graph that uses the variables for the axes. The slope of the objective function is then indicated, and the feasible solution having the optimal objective function value is determined by moving the objective function line toward or away from the origin.

Having said this, let us now develop together the constraints and the objective function for a specific example.

Assume that we have been monitoring truck production in a small emerging nation. The rate at which trucks are produced is limited primarily by the number of man-hours available from skilled welders and the number of kilowatt hours of electricity that can be supplied to the plant. Currently, the countries only active plant is making two models of a 10-ton truck, the Edsel and the Mack. The Edsel requires 30 manhours of welder time per truck and 20 kilowatt-hours of electricity. The Mack requires 40 manhours of welding time, but only 5 kilowatt-hours of electricity per truck. What is the maximum number of trucks that the plant could produce per day, if it is known that normally the plant has available 100 kilowatt-hours of electricity and 410 manhours of welder time per day?

To set this problem up we first need to decide what our unknowns (variables) will be. Since we are interested in the number of trucks produced per day, let's try letting our variables represent the number of each type of truck produced day:

X1 = Number of Edsels produced per day

X2 = Number of Macks produced per day

ANALYSIS ACTIVITIES

Assessing Constrained Systems

Having defined the variables we now can proceed to develop the constraints and the objective function. Let's do the objective function first. Our goal is to determine the maximum number of trucks that can be produced per day. This can be represented mathematically simply by summing the two variables that we defined: X1 + X2. But we want to maximize the value of this sum, so we'll set the sum equal to some variable (we'll use Z) and solve for the maximum possible value of Z. (Other problems, minimization problems, may require us to solve for the minimum possible value of Z to obtain an optimal solution.)

$$Z = X1 + X2$$

This is our objective function.

The constraints are, like the objective function, mathematical expressions. They set specific limits on the values that the variables (X1 and X2 in this case) can assume. For instance, we can not make a negative number of trucks per day, so neither X1 nor X2 can be less than zero.

$$X1 \geq 0$$

$$X2 \geq 0$$

In general, the variables used in linear programs cannot assume negative values, and <u>Nonnegativity Constraints</u> generally are assumed to exist unless otherwise specified. Most computer programs make this assumption (although it usually can be altered if necessary).[38]

Additional constraints can also be specified for our problem. We know that we cannot use more than 100 kilowatt-hours of electricity per day, but how can we relate this to the number of trucks that are produced?

Think back to the definitions of our variables, X1 and X2. They are specified in units of trucks per day. Since we know the amount of electricity available in kilowatt-hours per day, we can multiply each variable by the number of kilowatt-hours it uses per truck (per Edsel for X1) to determine the total electricity used per truck type. Summing the amount used for both truck types will give us thie total electricity used (which we know cannot exceed 100 kilowatt-hours).

----

38 Note that either of these nonnegativity constraints could have been made superfluous if for political or other reasons management had decreed that at least 5 of each type of truck must be made each day. In this case not only are X1 and X2 prevented from going negative, they cannot be smaller than 5. Note that this additional limitation may cause the final "permitted" solution to be less optimal than we could otherwise have determined, and may even prevent us from finding any feasible solution at all.

PROBLEM SOLVING TECHNIQUES
Assessing Constrained Systems

Our first constraint therefore becomes:

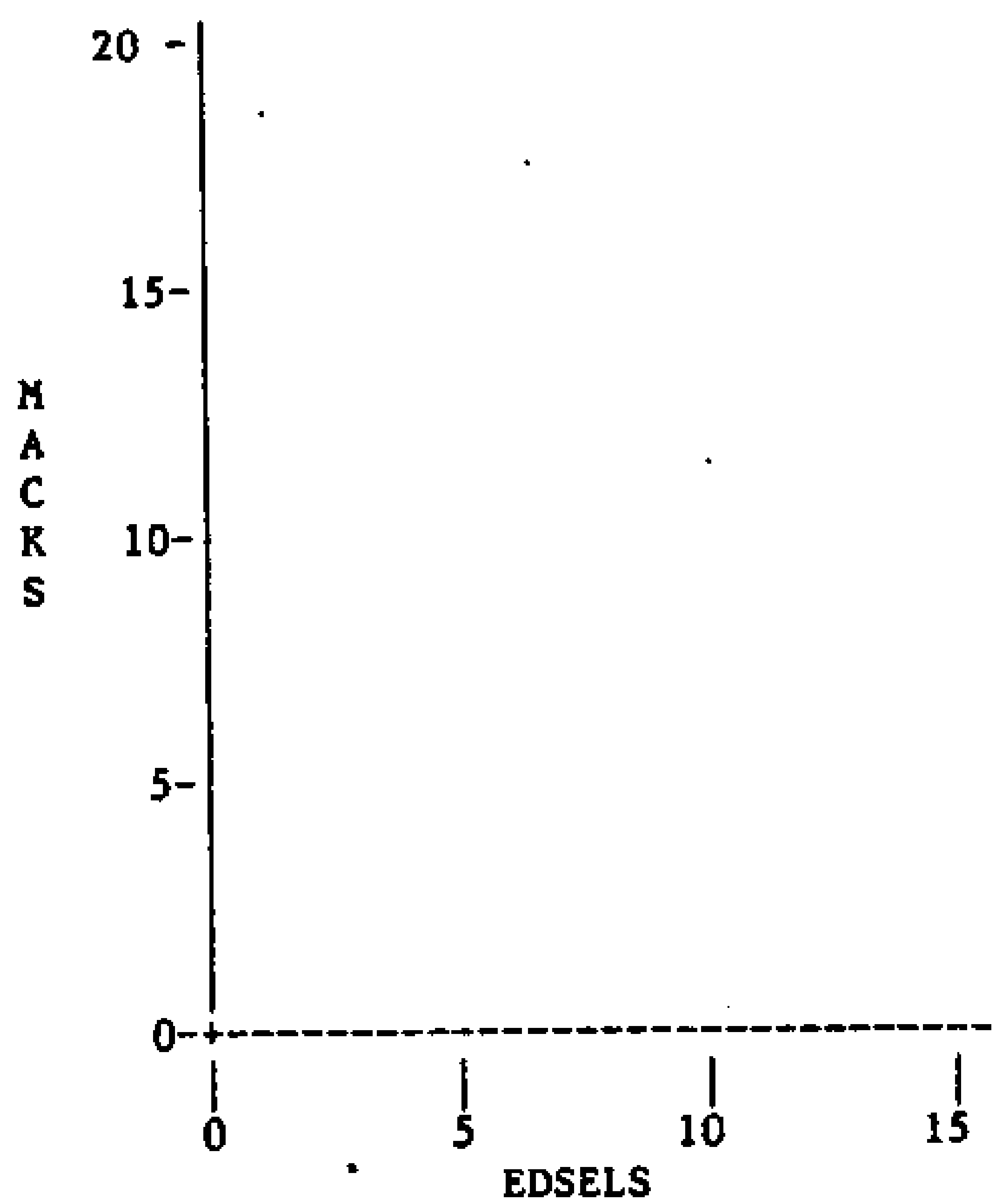
$$20X1 + 5X2 \leq 100$$

Another constraint can be established in a similar manner for the number of welding manhours that are used. See if you can specify this constraint on your own.

You should have obtained a constraint equation resembling the following:

$$30X1 + 40X2 \leq 410$$

We now have completed developing the mathematical expressions that comprise our linear program. So how do we solve it to determine the optimal solution? We'll try the graphical approach described above. First, use the following graph to plot the 1st quadrant portion of the lines described by changing our constraint inequalities into equalities.



ANALYSIS ACTIVITIES

Assessing Constrained Systems

Each of the lines you have plotted on the graph divides the two dimensional space of the graph into two distinct regions. On one side of each line are coordinates for X1 and X2 that do not satisfy the constraint inequality used to specify the line. The other side of the line (and the points on the line itself) represent feasible solutions to the constraint. To determine which side of the line the feasible solutions lie on, simply substitute values of any (X1,X2) coordinate not located on the line into the original constraint inequality. If the inequality is true, you're on the feasible solution side of the line. For example, let's use the origin as our trail set of coordinates. For the electricity constraint:

$$20X1 + 5X2 \le 100$$

becomes

$$(20 * 0) + (5 * 0) \le 100$$

$$(0) + (0) \le 100$$

$$0 \le 100$$ Is TRUE

Therefore the origin is on the "solution side" of the line. The origin also falls on the solution side of the second constraint. Also, if you think about the nonnegativity constraints that we developed, you will see that the X1 axis and the X2 axis define these constraint lines. The coordinates lying in the first quadrant can be determined to lie on the solution side of both nonnegativity constraints.

To be a feasible solution to the linear program we developed, the solution coordinate must be on the solution side of all of the constraints simultaneously. Hence, we are left with a region in the graph (bounded by the X1 axis on the bottom, the X2 axis on the left, the electricity constraint on the right, and the welding constraint on top) that contains an infinite number of coordinates, all of which represent represent feasible solutions to the problem. But which is the optimal solution?

To determine the optimal solution we must make use of the objective function. Since we simply set our sum equal to a variable, rather than a constant, we cannot simply graph a line and look for an intercept. However, we can note that the objective function gives us enough information to determine that it belongs to a family of lines, all having identical slopes. The optimal solution will be that represented by the coordinate in the feasible solution region giving the largest value of Z in the objective function. Let's consider the family of objective function lines for a few moments. One of the lines passes through the origin (which is a feasible solution). The value of Z for this coordinate (the origin) is X1 + X2 = 0 + 0 = 0. A pretty small value since neither X1 nor X2 are permitted to be negative.

Note that as we determine the value of Z for other lines passing through the feasible region, the value of Z continues to increase as we get farther and farther from the origin. Can you see that the maximum value of Z will be produced by the last point (coordinate) of the feasible region that is intersected by one of the objective function lines as we move away from the origin. In this case, that last point is the one defined by the intersection of the two constraint equations: (3,8). Hence, the maximum number of trucks that can be made per day is 11.

$$X1 + X2 = 3 + 8 = 11$$

Does this mean that the plant will produce 11 trucks per day? NO! At least not necessarily. It means that it has the capability to do so; but as an analyst you must still decide whether or not the intent is present, and whether or not the plant's decision makers are likely to be able to optimize their production effort. What the figure of 11 trucks per day represents, is simply an upper limit on the production capability of the plant. That limit may or may not be met, but it will not be exceeded (unless, of course, the data you used to develop the constraints was faulty). If your constraint data are in fact "soft," you had better do some sensitivity testing to estimate likely margins of error.

MULTIVARIABLE PROBLEMS

Multivariable problems are set up in the same way that two dimensional problems are; but, except for difficult three dimensional attempts, graphical solutions are not possible. The Simplex Method, usually done in a computer, is thus the way to go. A good discussion and illustration of the Simplex Method (including how to do it by hand) is contained in An Introduction to Management Science, Quantitative Approaches to Decision Making, Anderson, D.R., Sweeney, D.J., and Williams, T.A., West Publishing Company, New York, 1979, pages 73-124.

ANALYSIS ACTIVITIES
Assessing Constrained Systems

GOAL PROGRAMMING

BACKGROUND

Goal programming differs from linear programming in that it provides a methodology for dealing with multiple goals. Linear programming yields an optimum value in terms of one goal, i.e. the objective function. In reality, it is very difficult to reduce a problem down to a single objective or goal. Most problems we attempt to deal with are very complex and often include not only multiple goals but often even conflicting multiple goals. Goal programming is a technique developed in response to this multiple-goal problem.

Goal programming was developed by A. Charnes and W. Cooper as a result of their research on linear programming. Their developments were first published in 1961. Applications of goal programming have been scarce. Most of the recent treatment in the operations research literature has been by Sang M. Lee. Its slow development is due to its heavy mathematical orientation. However, with the increasing availability of computers and appropriate software, this should become an increasingly important methodology.

There are three main steps in performing a GP analysis:

1) Decide the resources available (i.e. the constraint equations)

2) Determine and rank goals (i.e. the objective functions)

3) Determine the degree of goal attainment

ADVANTAGES OF GOAL PROGRAMMING

1) Flexibility

2) Forces decision maker to think of goals and constraints in terms of their importance to the problem.

3) Ability to deal with multiple variables and multiple goals.

PROBLEM SOLVING TECHNIQUES
Assessing Constrained Systems

LIMITATIONS OF GOAL PROGRAMMING

1) Same as with all models, it is only as good as the input that is used.

2) Relationships must be linear or near linear.

3) Additivity: The constraints must be expressed in like units, otherwise the sum of the deviations from each goal will not yield a meaningful total deviation.

4) Divisibility: Can yield a fractional solution that is not always meaningful, i.e. 33.3 people.

5) Deterministic: All coefficients must be constants (no mechanism for dynamics).

Two variable goal programming problems can be solved graphically. As with linear programming, multivariable problems must be solved with a variation of the simplex method. This can be done by hand, but is extremely cumbersome and is best left to the computer.

NETWORK ANALYSIS

INTRODUCTION

Networks are interconnected paths over which things move. The things can be such things as automobiles (in which case we are dealing with a network of roads), oil (with a pipeline system), time (with projects being conducted concurrently or in sequence), electricity (with wiring diagrams or circuits), information signals (with communication systems), or people (with elevators or hallways).

In dealing with networks, we frequently are concerned with such things as maximum throughput of the system, the shortest (or cheapest) route between two or more locations, or with choke points or bottlenecks in the system. In this section we shall deal with a graphical method for assessing maximal throughput. Calculations on choke points will be dealt with in the section on queuing (waiting-line theory).

TERMINOLOGY

Arc
: Arcs are the individual paths (e.g., roads, rail lines, pipes) that make up the network.

Node
: Nodes are the places where arcs interconnect (e.g., rail switching yards, road intersections, electrical fuse boxes).

Sink
: A place (often a theoretical construct) where material flowing through the network ends up (appliances in a home for electricity disappearing from an electrical network, tankers for oil leaving a pipeline).

Source
: A place (often a theoretical construct) from whence material entering a system is thought to come (oil wells for a pipeline system, or Route 66 for Arlington traffic).

PROBLEM SOLVING TECHNIQUES
Assessing Constrained Systems

MINIMUM CUT-MAXIMUM FLOW PRINCIPLE

The Minimum Cut-Maximum Flow principle states that for any network with a single source and a single sink, the maximum feasible flow from source to sink equals the minimum cut value for any of the cuts through the network. Cuts are merely lines drawn across the network (perpendicular to the direction of flow through the network). The value of any cut line equals the sum of the capacities of the arcs it intersects.

Cut Method

1. Draw a topologic map of the network.
2. Indicate the flow capacity of each arc.
3. Draw all cut lines that might have a small value.
4. Determine the cut values.
5. Find the smallest cut value.

ANALYSIS ACTIVITIES
Assessing Constrained Systems

QUEUING THEORY

TERMINOLOGY

Queuing Theory
The operations research term for the study of waiting lines.

Queue
A waiting line. Also used to refer to the length (time or number) of the line.

Mean Arrival Rate
The average number of units arriving (entering the queue) per time period.

Mean Arrival Time
The average expected time between arrivals.

Mean Service Rate
The expected average number of units that can be serviced by one service station per time period at maximum service efficiency (no waiting for units to arrive -- no dead time).

Steady State
The normal operating characteristics of the waiting line after an initial start-up or transient period.

Utilization Factor
The proportion of time that the the service facility (ies) are in use. (This happens to equal the mean arrival rate divided by the mean service rate.)

GENERAL CASES

We all have experienced waiting in lines, whether at the grocery store or at college registration times. (Some buildings, I hear, even have lines waiting for elevators.) The interesting thing about lines, from an analytical viewpoint, is that a fair amount of work has been done determining methods to describe and to predict the operating characteristics of many waiting line situations. Thus, if we know something about the way that a given service station performs, we can estimate many items of interest about both the system itself and the units that it serves:

PROBLEM SOLVING TECHNIQUES

Assessing Constrained Systems

1. The percent of time or probability that the service facilities are idle
2. The probability of a specific number of units in the system
3. The average number of units in the system
4. The average time each unit spends in the system (service time plus time spent waiting)
5. The average number of units waiting for service
6. The probability that an arriving unit will have to wait.

In general, to develop a mathematical description of a service system (such as ferry boat crossing or a railroad switching yard) we need to know:

1. The arrival distribution (the histogram describing the relative likelihood of specific time intervals between consecutive arrivals).
2. The service time distribution (the probability density function describing the relative likelihood of specific service times).
3. The waiting line discipline (the rules governing which unit gets served next -- e.g., first come, first serve; or some sort of priority system, like toughest guy first or women and children first).

Fortunately, enough studies have been done that even if we do not have exact data, relatively good guesses can be made about the first two of these items if we know something about the topology (layout) of the system and about its discipline. Therefore, even if the Soviets won't let us make detailed measurements on the operation of its message centers, we still can calculate some reasonably good estimates of their performance capabilities if we can find out something about their layout (number of incoming and outgoing lines), queue discipline, mean message arrival time, and mean service time per message. Conversely, if we know something of a system's output capabilities, we can do some reverse engineering to estimate some of these parameters that we do not know.

Equations have been derived for many standard types of waiting line disciplines: single line-single service station, single line-multiple service stations, etc.-- for differing probability density functions for the arrival and service times in each case. For extremely complicated queuing systems, (e.g., personnel systems of large agencies) simulation models often are necessary to make the queuing model responsive to user needs.

A SPECIFIC EXAMPLE

For the specific case of a single channel waiting line that has Poisson arrivals and Exponential service times, and a first-in-first-out queue discipline, the system characteristics can be described in terms of two parameters:

$\bar{A}$ = the mean arrival rate

$\bar{S}$ = the mean service rate

Under these conditions:

The probability that the service facility will be idle is:[39]

$$P<o> = 1 - \frac{\bar{A}}{\bar{S}}$$

The probability of "n" units in the system is:

$$P<n> = \left|\frac{\bar{A}}{\bar{S}}\right|^{n} * P<o>$$

The average number of units in the system is:

$$L = \frac{\bar{A}}{\bar{S} - \bar{A}}$$

The average time a unit spends in the system (waiting time + service time) is:

$$W = \frac{L}{\bar{A}}$$

The average number of units in the queue is:

$$L<q> = \frac{\bar{A}^2}{\bar{S}\ (\bar{S}-\bar{A})}$$

The average time a unit spends in the queue is:

---

[39] The word processing device used to print this manual does not have the capability to subscript. The symbols enclosed by the "<" and ">" brackets normally are written as subscripts. P<o> would normally be a capital P with a subscripted zero and be called "P Nought" or "P sub-zero."

$$W\langle q\rangle = \frac{L\langle q\rangle}{\bar{A}}$$

The probability that an arriving unit will have to wait is:

$$P\langle w\rangle = \frac{\bar{A}}{\bar{S}}$$

MODELING AND SIMULATION

TYPES OF MODELS

In general we may define a model as a replica, or representation, of an idea or of an actual system (either physical or abstract). We don't think of words and numbers as models, but they are. They represent ideas. A picture, a map, a globe, a calendar, and a clock are models also. The first three represent something physical and the latter two represent time.

Models may be either descriptive or normative. A descriptive model is one that represents the nature of something but does not indicate any course of action that should be taken.

A normative model may contain some descriptive segments, but it differs from the descriptive model in that it is possible for a best, or preferable, ccurse of action to be indicated. It is prescriptive. This implies that some objective is included in the model and that it is possible to trace the effect of an action on that objective. We refer to the "best" action as the one that gives maximum satisfaction of the stated objective and/or minimizes the expected regret or loss associated with the action.

Models can also be classified as physical or abstract. A physical model, although representative of something else, has physical properties of its own. A globe and a model airplane are examples. An abstract model is not a physical item itself, although the item it represents may be physical rather than abstract. Mathematical models are, of course, abstractions.

The remaining classifications that will be discussed are, basically, further classifications of abstract models. The first of these classifications is that of deterministic versus stochastic (uncertainty) models. In a deterministic model the relationships are completely known and specified explicitly A model may be completely deterministic, or it may have stochastic characteristics also. It is convention to refer to a model that has any uncertainty incorporated into it as a stochastic model, even though it also has deterministic properties.

Occasionally we may decide to use a deterministic model in a situation that has uncertainty. When we do, we are implicitly assuming that the variance we are ignoring, or assuming away, is immaterial to the results. Whether or not this is justified depends on the situation. The loss of accuracy in the results may well be offset by the ease of constructing and understanding the model.

Another classification scheme is linear versus nonlinear models. Linear models are those that use only linear equations to describe the relationships. It is not necessary that the situation itself be linear, only that it be capable of description by linear equations. In contrast, nonlinear models utilize any type of mathematical function. Because nonlinear models are more difficult to work with and are not always capable of being analyzed, we may be able to justify some compromises in order to use a linear model. However, a linear model may not always be an acceptable way of reflecting the situation in question.

Our next classification involves static and dynamic models. A static model assumes that there is a given time period and that a fixed state of nature is in existence for that time period. Then, without regard to previous or future periods, an optimal course of action is sought. A dynamic model, on the other hand, considers several time periods and does not ignore the effect that an action in period 1, for example, might have on period 2. If the states of nature do not chance (are not expected to chance) from one time period to the next, then a static model can be used safely because the optimal strategy would be the same for each period.

Finally, we can classify models into those that are solvable and those that must be simulated. A solvable model is one in which there is an analytical way of finding a unique optimal course of action. Many problems require such a complicated set of equations to describe them that there is no way to analyze the model for an optimal action. In such cases we can resort to simulation. Rather than seeking the optimal solution, simulation requires the user to propose a set of possible solutions. These proposals are then introduced into the model, which typically is coded on a computer and verified for feasibility. If a proposal is feasible, the effect on the objective is projected. From the courses of action chosen to be evaluated we can select the one with the best result. Of course, it is possible that the best alternative was never included in the set to be evaluated. Simulation is very much like an experimental laboratory.

OVERVIEW

A model is merely a simplified representation of some part of reality. It is simplified because reality is too complex to copy exactly and because, for the modeler's purposes, much of the complexity is irrelevant

We have all build at least two types of models: physical and subjective. As hobbyists most of us have built physical models, models of planes, ships, or automobiles. (Design engineers build this kind of model to develop and test their products.)

As analysts, all of us have built subjective models. At first our models are purely inductive; we wade through detail and draw conclusions from decision-making structures, a total picture--a gestalt or synthesis--emerges. In other words, the analyst begins to have a gut feeling about situations and events. At this point, he has developed an intuitive subjective model of his country that better prepares him to cope with the masses of data and to understand why events have occurred and what can be expected to happen in the future.

ANALYSIS ACTIVITIES

Modeling and Simulation

Since all of us walk around with subjective models in our heads, the question is no longer whether to model, but how best to model.

The mental model is fuzzy, incomplete, and imprecisely stated. Furthermore, within one individual, a mental model quickly changes, sometimes even during the flow of a single conversation. The human mind assembles a few relationships to fit the context of a discussion and as the subject shifts so does the model. When only a single topic is being discussed, each participant in a conversation often is found to have employed a different mental model to interpret the subject. Fundamental assumptions differ and are never brought into the open. Goals differ and are left unstated.

A model, on the other hand, is a powerful conceptual device that can advance reason over rhetoric in the determination cf policy. A model is not, as sometimes supposed, a perfectly accurate representation of reality that can be trusted to make better decisions than people. It is a flexible tool that forces the people who use it to think hard and to confront one another, their common problems, and themselves directly and factually.

A formalized model differs only in complexity, precision, and explicitness from the informal subjective explanation or "mental model" that people ordinarily construct to guide them toward a goal. It is an account of the total set of forces that are believed to have caused and to sustain some problematic state of affairs. Like the informal mental model, they are derived frcm a variety of data sources, including facts, theories, and educated guesses. Unlike the mental model, they can be comprehensive, unambiguous, flexible, and subject to rigorous logical manipulation and testing.

When one develops a formal mathematical model, one gains the ability to determine the dynnmic consequences when the assumptions within the model interact with one another. The human mind is not well adapted to deciphering the consequences of a mental mode. Even when the mental model may be correct in structure and assumptions, the human mind--either individually or as a group consensus--is apt to draw the wrong conclusions. There is no doubt a computer can routinely and accurately trace through the statements of behavior for individual points in the model system.

The inability of the human mind to project accurately the consequences of its own mental models is clearly shown when a computer model is constructed to reproduce the assumptions held by a single person.

If a model is refined until it is fully agreeable in all its assumptions to the perceptions and ideas of a particular person, in many cases the model will not respond the way the person had anticipated. Ordinally, one's assumptions about structure and internal motivations are more nearly correct than are the assumptions about the implied behavior.

The flexibility of a computer model is one of its most important virtues. If you and I disagree about some aspect of the causal structure of a problem, we can, usually in a matter of minutes, run the model twice and observe its behavior under each set of assumptions. We may well discover that our argument was trifling, since the matter of interest to us is not changed by the altered assumptions.

A computer model constructed and used by an analysis or policy-making group has these advantages:

--It requires the analysts to improve and complete the rough mental sketch of the causes of the problem that they have in their heads.

--In the process of formal model-building, the builders discover and resolve various self-contradictions and ambiguities among their implicit assumptions about the problem.

--Once the model is running even in a rudimentary fashion, logical ''bootstrapping'' becomes possible. The consequences of promising but tentative formulations are tested. Observation of model behavior gives rise to new hypotheses about structure.

--If an acceptable standard of validity can be achieved, formal policy experiments can quickly reveal the probable outcomes of policy alternatives; novel problems may be discovered.

--Sensitivity analysis of the model reveals the model parameters having the greatest overall impact. If unknown, or poorly understood, which is generally the case in social planning, the ones that most affect model behavior can be investigated first (or given the most investigative effort).

--An operating model can be used to communicate concisely and precisely with people who were not involved in building the model.

Nevertheless, computer models present some potential pitfalls and difficulties. There are a wide variety of modeling approaches. For successful application, all require adequate knowledge and perception on the part of the persons formulating the specific model. An oversimplified model that fails to capture the essence of the system it presumes to represent, seems, superficially at least, to be a rigorous and complete analysis. In modeling, as with any other analytical technique, the final product is only as good as the expertise that lies behind its formulation and use.

CROSS-IMPACT ANALYSIS

This section has been adapted from a portion of the HANDBOOK OF FORECASTING TECHNIQUES, Part II, Description of 31 Techniques, prepared for the U.S. Army Engineer Institute for Water Resources, Fort Belvoir, VA . The work was done under Contract No. DACW31-75-C-0077, August 1977.

ABSTRACT

Cross-impact analysis strives to identify interactions among events or developments by specifying how one event will influence the likelihood, timing, and mode of impact of another event in a different but associated filed. Cross-impact analysis is used not only to probe primary and secondary effects of a specified event, but to improve forecasts and to generate single forecasts (or scenarios) from multiple forecasts. The method is also a tool useful in many sophisticated forecasting techniques.

DEFINITION

Cross-impact analysis is a systematic means of studying the interactions among events or developments. The underlying assumption is that a change in technology, social practices, values, or any other area will affect the surrounding environment in three ways: (1) it will change the probability of occurrence of interconnected events, (2) it will change the timing of interconnected events, and (3) it will affect the mode of impact of interconnected events.

HISTORY

The cross-impact method was developed during the 1960s by Gordon and Helmer as an outgrowth of objections that the Delphi technique of forecasting often failed to consider the interconnections among events. The method has found widespread use in technological as well as social forecasting. Perhaps the most advanced cross-impact analyses have been carried out at the Institute for the Future.

PROBLEM SOLVING TECHNIQUES

Modeling and Simulation

MAIN USES

The main use of the method is to describe and quantify the impact of one development upon others. Most frequently it is used to explore the implications of an advance in technology. Cross-impact analysis is often used to compare the implications of differing forecasts in a given field and to combine separate forecasts in discrete but allied fields into a single forecast.

Results have been used to define primary and secondary effects of a change (or a forecasted change) in one area upon: social trends, new or altered social demands, institutional functions (personnel, production, distribution, etc.), development in allied technologies, markets, long range plans, management policies, R&D decisions, and similar areas. The technique provides insight into trade-off options and is useful in testing the consequences of various policy actions. By clarifying the critical events underlying possible future developments, the method helps to identify the contingencies of future profit potentials for proposed programs. Computerized runs of cross-impact matrices establish better estimates of the probability of occurrence of individual events, facilitate tests of the most effective responses under various circumstances, and point to the key events to be monitored for future planning.

LIMITS AND CAUTIONS

Cross-impact analyses cannot allow for events not included in the matrix. Therefore defective or inadequate models of interacting elements can yield misleading results. Two other problems are evident. First, assignment of probabilities of interactions are subject to much uncertainty, especially in areas whose origin and anatomy are not clear. Examples include such "soft" trends as alienation, changing priorities and values, political disruption, or crime. Second, as pointed out by Bright, "use of one forecast at a time as 'true' is clearly erroneous, since the future results from the interactions of a number of events, some simultaneous and others very early. These events, coming in clusters, often must create a very different impact on other future events than the impact from events taken individually."

OTHER TECHNIQUES

Intuitively, participants in brainstorming, scenario writing, and Delphi forecasting make use of cross-impact analysis, although not necessarily in a focused or systematic fashion. The technique is used as part of such forecasting methods as feedback analysis and relevance trees. It is a central element in interactive simulation techniques of of forecasting, such as Forrester type dynamic models or KSIM.

ANALYSIS ACTIVITIES

Modeling and Simulation

PROCEDURES

In the usual procedure the first step is to construct a matrix consisting of the events of items to be cross-impacted. This can be done in a variety cf ways. As symmetrical matrix with identical items along both axes is perhaps most commonly used. The specific event or development is then assumed to occur. Its nature, mode, timing, etc., are specified. One then specifically inquires what the impacts of this development would be on all the other items of the matrix. At a minimum estimates are made of the influence of the first event on the probability, timing, and impact of the other items in the matrix. In some cases mini-scenarios of the interactions between two items are prepared. Usually, experts are used to make the cross-impact assessments. Simple matrices can be done by hand; elaborate matrices require computerization.

PRODUCTS OR RESULTS

The core output is a table or series of tables showing interactions among items. Major studies may include numerous computer runs showing one-to-one interactions using a variety of assumptions; primary, secondary, and tertiary levels of impact; interactions among clusters of events (aggregated forecasts) or among subdivided events (decomposed forecasts). Very often interactions identified on a matrix are strung together on a time basis to create a scenario.

Results can be qualitative, quantitative, or both. They can be gross or detailed and fairly certain or highly speculative. They can cover any time span. Credibility of results will range from high to doubtful depending on the qualifications of the judges and the nature of the material being cross-impacted.

RESOURCES

The most crucial resource is imaginative people knowledgeable in the field begin cross-impacted. This is a difficult combination to find in many areas of engineering and science so it is often wise to leaven the judgment of experts with that of people of sound and imaginative judgment although not expert in the field.

Cross-impacting is difficult and demanding work. Computerized assistance is essential if complex simultaneous interaction must be considered.

Requirements of time and money depend wholly on the nature of the problem.

COMMENT

Cross-impact analysis is a basic forecasting tool helpful if not essential in most sophisticated forecasting. Every forecaster dealing with interacting trends should have an appreciation of the principles involved.

MORPHOLOGICAL ANALYSIS

This section has been adapted from a portion of the HANDBOOK OF FORECASTING TECHNIQUES, Part II, Description of 31 Techniques, prepared for the U.S. Army Engineer Institute for Water Resources, Fort Belvoir, VA. The work was done under Contract No. DACW31-75-C-0077, August 1977.

ABSTRACT

Morphological analysis is designed to examine all possible solutions to a specified problem by juxtaposing the elements of the situation in every combination. It is, in a sense, a comprehensive and systematic type of cross-impact analysis. The approach has been used most in solving technological problems. Recently, however, the method has proved excellent for projecting alternative futures for a nation, industry, region, area, etc., under a variety of explicit assumptions. It is not primarily a method of forecasting, but is superior to any other method for examining "fans of futures," the comparative likelihood that they will come to pass, and the specific assumptions that underlie their existence.

DEFINITION

Morphological analysis is a technique for systematically exploring all possibilities within a system. The basic principle involves identification of the sectors of a system (e.g., for an internal combustion engine, sectors would include fuel, materials of construction, operating atmosphere, etc.) and all factors or variants within each sector (e.g., factors within the fuel sector would include wood, gasoline, gas, oil, coal, electricity, nuclear energy, etc.). Sectors and factors represent the morphology or taxonomy of the system. These morphological components are next arrayed in a matrix and all possible combinations are considered (e.g., wood fuel with all possible operative atmospheres, etc.; this is repeated for all other fuels, etc.) The purpose, of course, is to uncover useful workable combinations that might be overlooked in a less systematic approach.

PROBLEM SOLVING TECHNIQUES

Modeling and Simulation

HISTORY

In recent times the morphological method was developed by the late Fritz Zwicky, a Swiss astronomer who taught at California Institute of Technology and did much of his work at Mount Wilson and Mount Palomar. Zwicky used morphological analysis as long ago as 1942 when he was temporarily engaged in rocket research. In 1961 he established a society for Morphological Research in Pasadena with himself as president. Zwicky has published a number of books describing the method.

Although formalized by Zwicky, the method is actually ancient. A mystic Majorcan monk named Roman Lull devised a "Great Art" diagram in the late 1200s which employed the technique to solve all problems of philosophy and metaphysics. Lull's work was handicapped by lack of a computer. Descartes and Leibnitz consciously used Lull's technique, as did the Jesuit Kircher, inventor1 of the magic lantern.

Its use in generating a systematic array of alternative futures--that is, its use in social forecasting--dates from the late 1960s when a group at Stanford Research Institute used the method to develop what they called the "field anomaly relaxation method" (FARM) of projecting alternative futures.

MAIN USES

Morphological analysis is used when comprehensive coverage of an area is desired. It is especially helpful in much-worked areas (like internal combustion engines) because it forces consideration of combinations of elements in a non-traditional context.

In social forecasting, the method is used for generating scenarios of probable, plausible, possible, and very unlikely futures. It is better adapted to projecting possible lines of development than to forecasting per se. It is a method for exploring possibilities, not for forecasting what will happen. A particular virtue of the method is that it facilitates exploration of societal consequences of specific assumptions (e.g., what happens to a region with and without a specified type of waterway). Basically, the technique is a macro cross-impact device since it specifically requires consideration of interactions among sectors and factors. It can be applied in the same ways cross-impact analysis is applied.

In other fields the method has been used to seek out fresh solutions to problems in jet engine design, astronomical problems, transportation systems, warhead developments, communication systems, and even "planetary engineering."

LIMITS AND CAUTIONS

The successful use of morphological analysis requires that its user be able to judge what "new" combination of factors are worth probing more deeply. Clearly this is a severe limitation because the essence of blocks to many new ideas is incorrect acceptance of the idea that "it won't work." Similarly, the scenario writer may incorrectly discard a line of social evolution on grounds that "it doesn't make sense."

Leibnitz and Descartes were concerned that morphological approaches might lead to mechanization of thought. Jonathan Swift expressed this limit when he described a contrivance invented by a savant of Laputa that randomly combines letters, in this way "permitting the most ignorant person to write books without the least help of genius or study."

A worse problem is that morphological matrices give rise to astronomical numbers of combinations. A simple morphological box of eight sectors each with four factors gives rise to 65,536 different combinations. If each combination requires half an hour to assess for feasibility, something like 15 man years is needed. Therefore, as a practical matter, the too-great-richness must be kept in bounds by instructing the computer not even to consider combinations defined as clearly impossible.

OTHER TECHNIQUES

Cross-impact analysis is the principal competing technique. Indeed, comprehensive and systematic cross-impact analysis is, in fact, morphological analysis. Divergence Mapping (described elsewhere) employs morphological analysis.

PROCEDURES

In The Morphology of Propulsive Power Zwicky describes the procedures of his method in the following terms:

1. "An exact statement is made of the problem which is to be solved. For instance, we may wish to study the morphological character of all modes of motion, or of all possible propulsive power plants, telescopes, pumps communication, detection devices, and so on. If one specific device, method or system is asked for, the new method immediately generalizes the inquiry to all possible devices, methods, or systems which provide the answer to a more generalized request.

   It will be found that the task of formulating the initial statement or definition of the problem on hand is far more exacting than most investigators not acquainted with the new method are inclined to think. In fact, one is hard put to find in the existing literature satisfactory definitions even of well-known devices like pumps, stationary power plants, telescopes, and so on. The exact definition of apparently simple devices like injectors will be found to be a most difficult task and I doubt whether the combined common sense and sophisticated knowledge of any group of men would suffice to produce such a definition.

2. The exact statement of the problem to be solved, or the precise definition of the class of devices to be studied will reveal automatically the important characteristic parameters on which the solution of the problem depends. For instance, in the case of telescopes, some of these parameters are the location of the telescope (medium in which it is embedded). The nature of the aperture A, the recording device R, the nature of the changes to which the light is subjected from A to R, the motion of te telescope, the sequence of operations, etc. The second step thus involves the study of all of these significant parameters.

3. Each parameter pi will be found to possess a number of ki different independent irreducible values pi(1) pi(2) . . . pi(k).

For instance, the parameter motion" of a telescope may have the independent values p(1), p(2), p(3) = translation in three directions; p(4), p(5), p(6) = rotatory motion; p(7), p(8) . . . p(12) = oscillation in the first six motions, etc. These matrices are written in the following scheme:

p1(1), p1(2), . . . p1(k1)]

p2(1), p2(2), . . . p2(k2)]

. . . .

. . . .

pn(1), pn(2), . . . pn(kn)]

If one element is encircled in each matrix and all the circles are connected, every resulting chain of circles represents one possible solution off the original problem. The above scheme of matrices, if used to construct an n-dimensional space, leads to a morphological box. The analysis is complete if either one or no solution will be found in every drawer of the box.

It is exceedingly essential that up to this point no questions be asked as to what value one or the other solution may have. Such premature curiosity almost always defeats the unbiased application of the morphological method. However, once all of the solutions are found, one must know their relation to any given set of adopted performance values.

4. The determination of the performance values of all of the derived solutions represents the fourth major step in the morphological analysis.

   Lest one wishes to get lost in an enormous confusion of details, the performance evaluation must be carried out on a universal, although necessarily simplified basis. This is not always an easy task.

5. The final step involves the choice of particularly desirable special solutions and their realization.

   The conviction that all solutions can be realized is inherent in morphological thought. It may, of course, happen that some among the many solutions are of a relatively trivial nature."

PROBLEM SOLVING TECHNIQUES

Modeling and Simulation

SIMULATION MODELING

Simulation models are mathematical descriptions of the interrelationships believed to determine a system's behavior. They differ from other types of models in that the equations which comprise simulation models can not be solved simultaneously. This property is not actually a requirement or a by-product of simulation modeling techniques or computer languages, but rather reflects the reason one uses simulation modeling techniques. That is, one usually goes to simulation modeling when it is impossible or impractical to write and/or measure all of the variables necessary to solve the set of simultaneous equations that would fully describe a system.

Simulation may be used in either a stochastic or a deterministic situation. It most frequently is used when uncertainty is present. The reason is that when several stochastic processes are present in the same system, it frequently is difficult to describe analytically the net net result of their interaction.

In the deterministic case one must describe the decision situation with a set of equations. If the equations are beyond our analytical capability and we are unable to develop an adequate, simplified approximation, simulation can be used to find a reasonably good (approximate) solution. Many nonlinear problems are solved in this manner. The simulation model is used as a practical way of picking the best alternative from a set of proposed solutions, verifying feasibility, and comparing objective values.

A stochastic situation involves one or more random variables, i.e., variables described by some type of probability function. When a simulation used such probabilistic variables, the model sometimes is referred to as a "Monte Carlo" study, after the famous gambling capital of Monaco.

Simulation models work by incrementally estimating the degree to which a system changes over a period of time. They do so by calculating the current values of all system state variables and then determining the instantaneous rate at which these variables are changing at that point in time. These rates of change are then assumed to hold constant for a short period of time so that the new values of the state variables can be determined based on the the amount of change that was predicted from the assumed constant rates of change. The new values of the state variables are then used to recalculate the instantaneous rates of change, and the process is repeated (again and again) until the entire time period that was desired to be covered by the simulation has been covered.

Simulation models thus are designed to follow (predict) changes over time by making a series of sequential approximations. Each approximation is based (uses as its starting point) the final result of the preceeding approximation. Other modeling methods differ in manner in which the final results are achieved. Econometric models (about which we'll here more later), for instance, solve directly for the final answer using a complete set of simultaneous equations.

ANALYSIS ACTIVITIES

Modeling and Simulation

Because of their successive approximation approach, simulation models have the advantage of permitting the time-course of changes to be displayed. Thus, one is not limited to solely a before and after picture, but also can see what is happening during the simulation interval and how the system reached the final state.

In general, simulation models are most useful in making long-range forecasts where exact numerical estimates are not needed. That is, other modeling approaches usually are more accurate for short-term forecasting. But, because simulation models permit one (because of the successive approximation method) to switch the "operating rules" during the middle of a simulation period--if changes in the state variables suggest that the real-world would, under similar circumstances, operate under new rules--dramatic changes in trends, etc. are frequently predicted more accurately via simulation techniques. Still, simulation models usually are most effective when used to compare the impact of alternative scenarios (of policy decisions or natural phenomena, etc.).

Our coverage of simulation modeling in class will be at the overview level only, intended to give you a feel for what it is like to work with a completed model, and how such models should (and should not be) used.

INPUT-OUTPUT MODELING

Input-output analysis is a name given to a modeling procedure in which the output product of an industrial sector is set equal to the input consumption of that product by other industries and consumers. Unlike many other models, input-output models are often very disaggregated and can, therefore, show more casual relationships. Input-output models can be interpreted in terms of block diagrams and matrix algebra techniques. This modeling approach has been applied to a variety of economic policy analysis and forecasting problems. This section will develop some of the basic theories of input-output modeling.

The basic phenomenon described by an input-output model is that the inputs to a particular industrial sector are outputs from a number of other industrial sectors. Suppose there are four industries with known output rates. The output of each industry is used by itself, by the other three industries, and by consumers. The output rate of each industry can be equated to the sum of the input rates for that industry's product needed to supply each industry plus the required input rate for final demand to consumers. This relation can be used to determine the output rates required to yield any final demand from consumers.

There are several restrictions implied: no exports and imports (a restriction which can easily be removed), constant return to scale (which may not be a correct assumption), and no substitutability of products in the event of scarcity or high price (not always true in a competitive economy -- e.g., when the price of sugar rises, less sugar will be used in the output product). The usual matrix form of input-output models permits an industry to produce only a single output.

The total output, or output production, is used to satisfy interindustry demands and final demands. However, the activity of production also produces a number of outputs, many of them inadvertent, which have helpful or harmful effects on the industry producing the output, other industries, and society. These are externalities, or external effects. Production itself is internal to the producing industry and is generally considered in great detail by industrial management. In the past many externalities have been neglected, but with the increasing interaction of society and technology they can no longer be ignored. Input-Output models can expand to include many externalities, such as pollution of various types, education, or taxes.

Often the externalities will be caused not only by the total outputs but by the final demands as well. If the externality were taxes, for example, it might be computed as a function of total output (an industrial tax) and total demand (a sales tax on the consumer) as well. The externality may well be the demand function for another sector, and the required inputs from this sector may well contribute to the externality. If the externality were pollution, for example, then antipollution activities, such as different types of gasoline, may produce pollution externalities themselves. Further, the total outputs of the externality sector may well act to increase or decrease the final demands of the producers. There may also be some exogenous externalities introduced or removed, and these exogenous externalities may influence the final demand. We can proceed further and extend the input model to include both internal and exogenous variables. Some of the exogenous variables would represent total demand variables, and some would represent externalities. The general equations are simply a set of linear algebraic equations with constant coefficients.

These equations can be solved as long as the matrix representing their coefficients is nonsingular.** In addition, it is important that all the coefficients be nonnegative, since negative rates of production cannot exist. The above assumes that the coefficients are time-invariant and independent of the inputs and outputs of the system. This restriction can be removed; and there is motivation to do this. Production coefficients often depend upon production level, time, and possibly other variables.

** A matrix F is nonsingular if the determinant of F is nonzero and there exists a matrix G such that FG = GF = I. G is the inverse matrix of F.

ANALYSIS ACTIVITIES

Modeling and Simulation

ECONOMETRIC MODELING

OVERVIEW

Two quite disparate, but overlapping, opinions exist concerning input-output or cause-effect in public and societal systems. One opinion holds that the primary and causative factor in societal growth and decay is the change in the population level resulting from the various inputs to the rate variables in the population sector. Business and the economy is, in this theory, a function of population and natural resources. The alternative opinion assumes that population in a given sector responds to economic conditions. In this view economic conditions are input or causative factors and population the output or response due to economic conditions.

It probably would take a very special restricted society for one of these views to apply to the complete exclusion of the other. In a very traditional, closed society, where migration is very small, economic growth would appear to respond primarily to population inputs. In an advanced, rapidly changing society with bold, adventurous, and entrepreneurial type inhabitants, population growth and migration may well be directly responsive to economic climate. It is this second opinion of cause and effect which is emphasized in econometric models.

ECONOMETRICS AND ECONOMIC THEORY

The basic concern of economic theory is development of concepts which are useful in describing financial events. Econometrics is a quantitative approach to economic behavior in that it is concerned with measurements that provide empirical content for economic theory. Estimation of parameters and observation of data are the two primary tasks in econometrics.

Economic theory is used in econometrics to determine a set of functional relationships which mathematically characterize a problem. At least three relationships are generally involved: a demand function, a supply function, and an equilibrium or market clearing equation. Specification of the structure of the equations for an econometric model is the most critical step in that later statistical and validation procedures assume that the structure is correct. An incorrect or inappropriate structure can lead to misleading or incorrect interpretation of results from the model.

Specification of the structure of an econometric model is a statement of our *a priori* hypotheses concerning the functioning of an economic system. Specification of the structure completely determines our conception of the dependent and independent variables and their functional dependence upon one another. The demand function expresses the consumption quantity of an output product or service as a function of (1) its price, (2) other attributes of its price, (3) the price ·f complementary or supplementary goods cr services, (4) other attributes of complementary or supplementary goods or services, and (5) the economic and social environment of the product or service.

For example the demand for electricity for heating is a function of environmental variables such as average outside temperature and economic status of the people using the electricity. Demand for electricity for heating is also a function of the price of electric energy for heating, the price of competing forms of energy for heating (perhaps lagged by several years), and other attributes such as the cleanliness of electric energy for heating, compared with other energy forms, and initial costs of various types of heating systems.

It is important that we note that the demand is a function of the attributes of the product or service, including costs. That is, the demand for a product or service (or system) depends upon the nature of what is furnished. It is not fully meaningful to conduct a study of electrical needs in the United States over the next 20 years to determine demand for electricity without considering economic, social, and other attributes of the product (electricity) to be delivered. Environmental health hazards, costs, and alternative methods of achieving the same objective must always be considered

Once written, we can rearrange our demand equation, assuming it is well formed, to obtain an equation for the price people will pay for a given product or service in terms of demand, other product costs, and environment. By examining such equations we can determine that if we are considering a normal product or service that is wanted, the quantity desired, or demand, will increase as the price or cost of the product or service decreases.

The demand curve is not sufficient by itself to determine the price and quantity of a product or service. We also need a supply curve to describe the behavior of the supplier of the product or service (or system). This supply curve will represent the lowest price at which producers will supply a given output or the maximum quantity that will be supplied at a specified price. The demand curve will be monotone decreasing with increasing price, and the curve of price will be monotone decreasing with increasing demand.

Econometric practice assumes that the actual price paid for any product or service is determined by the intersection of the demand and supply curves (i.e.,supply equals demand). Suppliers of a product or service will continue to produce it until the minimum price is not met at a given level of production. Purchasers of products or services do not pay the maximum they are willing to pay, but rather the market price of a product.

ANALYSIS ACTIVITIES

Changes in any of the impacting variables in either the demand or the supply curve will produce a new market price. Prediction of the effects of these changes, requires a knowledge of the functional relationships involved in both equations. In practice the ratio of the rate of change in demand quantity to the rate of change in price is often computed. This ratio is known as price elasticity, and is dimensionless. We can approximately determine the effect on supplier revenue of a change in price from a knowledge of the elasticity. The change in supplier revenue is the difference between the new quantity and price and the old quantity and price, assuming the demand and supply equations are satisfied for both the old and the new situation. If the price elasticity is less than 1, we say that the demand function is inelastic in that a reduction in price will lower total revenue to the supplier, whereas if the demand function is elastic (>1), then a reduction in demand quantity by the consumer will result in a decrease in total revenue to the supplier.

Some products appear to be inelastic, whereas others appear elastic, at least over short enough time spans. Demand for a necessary commodity with a low normal price is generally inelastic -- drinking water, for example. Total revenue varies directly with the price charged for drinking water, unless a price increase were phenomenal. Demand for a discretionary commodity with a high normal price is usually elastic -- a backyard swimming pool or a recreational boat, for example. Knowledge of the elasticity, therefore, gives useful information concerning how revenue changes in a system as a result of changes in either price or demand.

IDENTIFICATION OF PARAMETERS

In order for solution of a system of simultaneous equations to be possible, there must be at least as many independent structural equations as there are independent (often called endogenous) variables. If, in constructing a model, we find that we have more endogenous variables than equations, we must somehow specify enough structural equations for there to be an equal number of independent structural equations and endogenous variables. The problem in econometric estimation and identification is to determine the interactions between all variables, although many of these interactions will be negligible or nonexistent. Often it will be desirable to associate dynamics (time dependent interactions), since physical, public, and societal systems are generally not in equilibrium over time. In many cases there are accuracy difficulties associated with least squares estimation. Several alternative approaches are available.

PROBLEM SOLVING TECHNIQUES

Modeling and Simulation

EXAMPLE OF A ECONOMETRIC MODEL

The following model is an unclassified model developed by the Office of Economic Research, CIA, using the TROLL software package.

The following are TROLL conventions:

* means multiplication

/ means division

Calculations proceed from left to right.
¢i.e., 10/5*2 is equal to 4 whereas 10/(5*2) is equal to 1]

cny(-1) is the value of cny in the immediately preceding period

cny(-2) is the value of cny two periods ago

MODEL: OILPRICE OPECLTSC

****

VERSION OF THE MODEL USED FOR INITIAL ANALYSIS OF ENERGY SUPPLY SCENARIOS FOR 1985 AND 1990.

****

SYMBOL DECLARATIONS

ENDOGENOUS:

```
   CNCPD    - CANADIAN CONSUMER PRICES
   CNED     - CANADIAN ENERGY DEMAND
   CNRATIO  - CANADIAN RATIO OF ENERGY DEMAND TO GNP
   CNRPE    - CANADIAN RELATIVE ENERGY PRICE
   CNY      - CANADIAN GNP
   EODH     - EXCESS OIL DEMAND:HIGH CASE (M/IE)
   EODL     - EXCESS OIL DEMAND:LOW CASE (M/IE)
   EODM     - EXCESS OIL DEMAND:MEDIUM CASE (M/IE)
   EODVL    - EXCESS OIL DEMAND: VERY LOW CASE (PSAC)
   JPCPD    - JAPANESE CONSUMER PRICES
   JPED     - JAPANESE ENERGY DEMAND
   JPRATIO  - JAPANESE RATIO OF ENERGY DEMAND TO GNP
   JPRPE    - JAPANESE RELATIVE ENERGY PRICE
   JPY      - JAPANESE GNP
   OECDED   - OECD ENERGY DEMAND
   OECDP    - OECD CONSUMER PRICES
   OECDRPE  - OECD RELATIVE PRICE OF ENERGY
   OECDY    - OECD GNP
   OSNOP  PCOPP  PTR  PTRS
```

RATIOO - RATIO OF OECD OIL DEMAND TO GNP
SNONOILP TOTVAL
USCPD - US CONSUMER PRICES
USED - US ENERGY DEMAND
USRATIO - US RATIO OF ENERGY DEMAND TO GNP
USRPE - US RELATIVE ENERGY PRICE
USY - US GNP
WECPD - WESTERN EUROPEAN CONSUMER PRICES
WEED - WESTERN EUROPEAN ENERGY DEMAND
WERATIO - WESTERN EUROPEAN RATIO OF ENERGY DEMAND TO GNP
WERPE - WESTERN EUROPEAN RELATIVE ENERGY PRICE
WEY - WESTERN EUROPEAN GNP

EXOGENOUS:
ADDADJ ADDDEC BETA
CNCPDB - CANADIAN CONSUMER PRICES-BASELINE
CNRPEB
CNYB - CANADIAN GNP-BASELINE
DEC DELTA DUM1 DUM2 EODLADJ
INCOME - INCOME ELASTICITY FOR ENERGY DEMAND -- VECTOR OF COEFFICIENTS
JPCPDB - JAPANESE CONSUMER PRICES-BASELINE
JPRPEB
JPYB - JAPANESE GNP-BASELINE
NOICC - NET OIL IMPORTS OF COMMUNIST COUNTRIES (M/IE)
NONODOS - NON-OECD NON-OPEC DOMESTIC OIL SUPPLY (M/IE)
NONOOD - NON-OECD NON-OPEC OIL DEMAND (M/IE)
NOSNOP
ODCOD - OTHER DEVELOPED COUNTRY OIL DEMAND
ODCOS - OTHER DEVELOPED COUNTRY OIL SUPPLIES
ODOD - OPEC DOMESTIC OIL DEMAND (M/IE)
ODOS - OECD DOMESTIC OIL SUPPLY (M/IE)
OECDSH OECDSL OECDSM OECDSVL OFF ON
ONOES - OECD NON-OIL ENERGY SUPPLY (M/IE)
OPXFPID - NOMINAL OPEC NOMINAL PRICE
OPXFPIDB - NOMINAL OPEC NOMINAL PRICE-BASELINE
POOSH POOSL POOSM POOSVL
PRICE - PRICE ELASTICITY FOR ENERGY DEMAND -- VECTOR OF COEFFICIENTS
RA RA1
USCPDB - US CONSUMER PRICES-BASELINE
USRPEB
USYB - US GNP-BASELINE
WECPDB - WESTERN EUROPEAN CONSUMER PRICES-BASELINE
WERPEB
WEYB - WESTERN EUROPEAN GNP-BASELINE

COEFFICIENT:
A1 - CONSTANT TERM IN US ENERGY DEMAND EQUATION
A2 - CONSTANT TERM IN WE ENERGY DEMAND EQUATION
A3 - CONSTANT TERM IN JP ENERGY DEMAND EQUATION
A4 - CONSTANT TERM IN CN ENERGY DEMAND EQUATION
YW1 - LAG-WEIGHT FOR GNP(T-0) IN ENERGY DEMAND EQUATION

```
YW2     - LAG-WEIGHT FOR GNP(T-1) IN ENERGY DEMAND EQUATION
YW3     - LAG-WEIGHT FOR GNP(T-2) IN ENERGY DEMAND EQUATION
```

EQUATIONS

```
     **************************

     CANADA:

     **************************

     ENERGY DEMAND (ED) AS A FUNCTION OF GNP (Y)
     AND DOMESTIC RELATIVE ENERGY PRICES (RPE):
  1:        LOG(CNED) = A4+PRICE*(0.4*LOG(CNRPE)+0.25*LOG(CNRPE(-1
            0.15*LOG(CNRPE(-2))+0.1*LOG(CNRPE(-3))+0.1*LOG(CNRPE(-
            )+INCOME*(YW1*LOG(CNY)+YW2*LOG(CNY(-1))+YW3*LOG(CNY(-2

     IMPACT OF HIGHER NOMINAL OPEC OIL PRICES (OPXFPID)
     ON REAL GNP (Y):
  2:        CNY = CNYB*EXP(0.004*LOG(OPXFPID/OPXFPIDB)+(-0.027)*LO
            OPXFPID(-1)/OPXFPIDB(-1))+(-0.021)*LOG(OPXFPID(-2)/
            OPXFPIDB(-2))+(-0.016)*LOG(OPXFPID(-3)/OPXFPIDB(-3)))

     IMPACT OF HIGHER NOMINAL OPEC OIL PRICES (OPXFPID)
     ON CONSUMER PRICES (CPD):
  3:        CNCPD = CNCPDB*EXP(0.042*LOG(OPXFPID/OPXFPIDB)+0.027*L
            OPXFPID(-1)/OPXFPIDB(-1))+0.019*LOG(OPXFPID(-2)/OPXFPI
            -2))+0.009*LOG(OPXFPID(-3)/OPXFPIDB(-3)))

     RATIO OF ENERGY DEMAND (ED) TO GNP (Y):
  4:        CNRATIO = CNED/3428/CNY*100

     **************************

     JAPAN:

     **************************

     ENERGY DEMAND (ED) AS A FUNCTION OF GNP (Y)
     AND DOMESTIC RELATIVE ENERGY PRICES (RPE):
  5:        LOG(JPED) = A3+PRICE*(0.4*LOG(JPRPE)+0.25*LOG(JPRPE(-1
            0.15*LOG(JPRPE(-2))+0.1*LOG(JPRPE(-3))+0.1*LOG(JPRPE(-
            )+INCOME*(YW1*LOG(JPY)+YW2*LOG(JPY(-1))+YW3*LOG(JPY(-2

     IMPACT OF HIGHER NOMINAL OPEC OIL PRICES (OPXFPID)
     ON REAL GNP (Y):
  6:        JPY = JPYB*EXP((-0.038)*LOG(OPXFPID/OPXFPIDB)+(-0.037)
            LOG(OPXFPID(-1)/OPXFPIDB(-1))+(-0.039)*LOG(OPXFPID(-2)
            OPXFPIDB(-2))+(-0.013)*LOG(OPXFPID(-3)/OPXFPIDB(-3)))

     IMPACT OF HIGHER NOMINAL OPEC OIL PRICES (OPXFPID)
     ON CONSUMER PRICES (CPD):
```

```
7:      JPCPD = JPCPDB*EXP(0.042*LOG(OPXFPID/OPXFPIDB)+0.062*L
        OPXFPID(-1)/OPXFPIDB(-1))+0.028*LOG(OPXFPID(-2)/OPXFPI
        -2))+0.01*LOG(OPXFPID(-3)/OPXFPIDB(-3)))

  RATIO OF ENERGY DEMAND (ED) TO GNP (Y):
8:      JPRATIO = JPED/5468/JPY*100

  **************************

  UNITED STATES:

  **************************

  ENERGY DEMAND (ED) AS A FUNCTION OF GNP (Y)
  AND DOMESTIC RELATIVE ENERGY PRICES (RPE):
9:      LOG(USED) = A1+PRICE*(0.4*LOG(USRPE)+0.25*LOG(USRPE(-1
        0.15*LOG(USRPE(-2))+0.1*LOG(USRPE(-3))+0.1*LOG(USRPE(-
        )+INCOME*(YW1*LOG(USY)+YW2*LOG(USY(-1))+YW3*LOG(USY(-2

  IMPACT OF HIGHER NOMINAL OPEC OIL PRICES (OPXFPID)
  ON REAL GNP (Y):
10:     USY = USYB*EXP((-0.014)*LOG(OPXFPID/OPXFPIDB)+(-0.014)
        LOG(OPXFPID(-1)/OPXFPIDB(-1))+(-0.014)*LOG(OPXFPID(-2)
        OPXFPIDB(-2))+(-0.007)*LOG(OPXFPID(-3)/OPXFPIDB(-3)))

  IMPACT OF HIGHER NOMINAL OPEC OIL PRICES (OPXFPID)
  ON CONSUMER PRICES (CPD):
11:     USCPD = USCPDB*EXP(0.023*LOG(OPXFPID/OPXFPIDB)+0.022*L
        OPXFPID(-1)/OPXFPIDB(-1))+0.011*LOG(OPXFPID(-2)/OPXFPI
        -2))+0.003*LOG(OPXFPID(-3)/OPXFPIDB(-3)))

  RATIO OF ENERGY DEMAND (ED) TO GNP (Y):
12:     USRATIO = USED/32722/USY*100

  **************************

  WESTERN EUROPE:

  **************************

  ENERGY DEMAND (ED) AS A FUNCTION OF GNP (Y)
  AND DOMESTIC RELATIVE ENERGY PRICES (RPE):
13:     LOG(WEED) = A2+PRICE*(0.4*LOG(WERPE)+0.25*LOG(WERPE(-1
        0.15*LOG(WERPE(-2))+0.1*LOG(WERPE(-3))+0.1*LOG(WERPE(-
        )+INCOME*(YW1*LOG(WEY)+YW2*LOG(WEY(-1))+YW3*LOG(WEY(-2

  IMPACT OF HIGHER NOMINAL OPEC OIL PRICES (OPXFPID)
  ON REAL GNP (Y):
14:     WEY = WEYB*EXP((-0.029)*LOG(OPXFPID/OPXFPIDB)+(-0.025)
        LOG(OPXFPID(-1)/OPXFPIDB(-1))+(-0.02)*LOG(OPXFPID(-2)/
        OPXFPIDB(-2))+(-0.01)*LOG(OPXFPID(-3)/OPXFPIDB(-3)))

  IMPACT OF HIGHER NOMINAL OPEC OIL PRICES (OPXFPID)
  ON CONSUMER PRICES (CPD):
```

```
15:     WECPD = WECPDB*EXP(0.025*LOG(OPXFPID/OPXFPIDB)+0.059*L
        OPXFPID(-1)/OPXFPIDB(-1))+0.017*LOG(OPXFPID(-2)/OPXFPI
        -2))+0.012*LOG(OPXFPID(-3)/OPXFPIDB(-3)))

  RATIO OF ENERGY DEMAND (ED) TO GNP (Y):
16:     WERATIO = WEED/21475/WEY*100

  *******************************

  MISCELLANEOUS RELATIONSHIPS

  *******************************

  EXCESS OIL DEMAND -- VERY LOW CASE:
17:     EODVL = DEC*(CNED+JPED+USED+WEED+ADDADJ+ADDDEC+EODLADJ
        ONOES+OECDSVL+ODOS)+(NONOOD+ODCOD-(NONODOS+ODCOS))+NOI
        ODOD-POOSVL

  EXCESS OIL DEMAND -- LOW CASE:
18:     EODL = DEC*(CNED+JPED+USED+WEED+ADDADJ+ADDDEC+EODLADJ)
        ONOES+OECDSL+ODOS)+(NONOOD+ODCOD-(NONODOS+ODCOS))+NOIC
        ODOD-POOSL

  EXCESS OIL DEMAND -- HIGH CASE:
19:     EODH = DEC*(CNED+JPED+USED+WEED+ADDADJ+ADDDEC+EODLADJ)
        ONOES+OECDSH+ODOS)+(NONOOD+ODCOD-(NONODOS+ODCOS))+NOIC
        ODOD-POOSH

  EXCESS OIL DEMAND -- MEDIUM CASE:
20:     EODM = DEC*(CNED+JPED+USED+WEED+ADDADJ+ADDDEC+EODLADJ)
        ONOES+OECDSM+ODOS)+(NONOOD+ODCOD-(NONODOS+ODCOS))+NOIC
        ODOD-POOSM

  SIMULATED OECD ENERGY DEMAND:
21:     OECDED = DEC*(CNED+JPED+USED+WEED+ADDADJ+ADDDEC)

  OECD GNP:
22:     OECDY = CNY*0.04+JPY*0.1+USY*0.48+WEY*0.38

  OECD CONSUMER PRICES:
23:     OECDP = CNCPD*0.04+JPCPD*0.1+USCPD*0.48+WECPD*0.38

  ***************************************

   OECD AGGREGATE RATIOS

  ***************************************

  RATIO OF OECD ENERGY DEMAND (OECDED) TO GNP:
24:     RATIOE = OECDED/63093/(CNY*0.04+JPY*0.1+USY*0.48+WEY*0
        )*100

  RATIO OF OECD OIL DEMAND TO GNP:
25:     RATICO = (OECDED-ONOES)/32618/(CNY*0.04+JPY*0.1+USY*0.
```

```
        WEY*0.38)*100

   RELATIVE PRICE OF ENERGY IN THE OECD:
26:        OECDRPE = 0.04*CNRPE+0.1*JPRPE+0.48*USRPE+0.38*WERPE

27:        CNRPE = DUM1*(CNRPEB*EXP(0.109*LOG(OPXFPID/OPXFPIDB)+(
           -0.012)*LOG(OPXFPID(-1)/OPXFPIDB(-1))+(-0.007)*LOG(
           OPXFPID(-2)/OPXFPIDB(-2))+(-0.003)*LOG(OPXFPID(-3)/
           OPXFPIDB(-3))))+DUM2*(CNRPEB*EXP(PTR*(0.7*LOG(OPXFPID/
           OPXFPIDB)+0.2*LOG(OPXFPID(-1)/OPXFPIDB(-1))+0.06*LOG(
           OPXFPID(-2)/OPXFPIDB(-2))+0.04*LOG(OPXFPID(-3)/OPXFPID
           -3))+0.*LOG(OPXFPID(-4)/OPXFPIDB(-4))+0.*LOG(OPXFPID(-
           OPXFPIDB(-5)))))

28:        JPRPE = DUM1*(JPRPEB*EXP(0.207*LOG(OPXFPID/OPXFPIDB)+
           0.194*LOG(OPXFPID(-1)/OPXFPIDB(-1))+0.023*LOG(OPXFPID(
           /OPXFPIDB(-2))-0.003*LOG(OPXFPID(-3)/OPXFPIDB(-3))))+D
           *(JPRPEB*EXP(PTR*(0.7*LOG(OPXFPID/OPXFPIDB)+0.2*LOG(
           OPXFPID(-1)/OPXFPIDB(-1))+0.06*LOG(OPXFPID(-2)/OPXFPID
           -2))+0.04*LOG(OPXFPID(-3)/OPXFPIDB(-3))+0.*LOG(OPXFPID
           )/OPXFPIDB(-4))+0.*LOG(OPXFPID(-5)/OPXFPIDB(-5)))))

29:        USRPE = DUM1*(USRPEB*EXP(0.192*LOG(OPXFPID/OPXFPIDB)+(
           -0.018)*LOG(OPXFPID(-1)/OPXFPIDB(-1))+0.001*LOG(OPXFPI
           -2)/OPXFPIDB(-2))+0.001*LOG(OPXFPID(-3)/OPXFPIDB(-3)))
           DUM2*(USRPEB*EXP(PTR*(0.7*LOG(OPXFPID/OPXFPIDB)+0.2*LO
           OPXFPID(-1)/OPXFPIDB(-1))+0.06*LOG(OPXFPID(-2)/OPXFPID
           -2))+0.04*LOG(OPXFPID(-3)/OPXFPIDB(-3))+0.*LOG(OPXFPID
           )/OPXFPIDB(-4))+0.*LOG(OPXFPID(-5)/OPXFPIDB(-5)))))

30:        WERPE = DUM1*(WERPEB*EXP(0.217*LOG(OPXFPID/OPXFPIDB)+
           0.025*LOG(OPXFPID(-1)/OPXFPIDB(-1))+0.01*LOG(OPXFPID(-
           OPXFPIDB(-2))+0.001*LOG(OPXFPID(-3)/OPXFPIDB(-3))))+DU
           (WERPEB*EXP(PTR*(0.7*LOG(OPXFPID/OPXFPIDB)+0.2*LOG(
           OPXFPID(-1)/OPXFPIDB(-1))+0.06*LOG(OPXFPID(-2)/OPXFPID
           -2))+0.04*LOG(OPXFPID(-3)/OPXFPIDB(-3))+0.*LOG(OPXFPID
           )/OPXFPIDB(-4))+0.*LOG(OPXFPID(-5)/OPXFPIDB(-5)))))

   ******************************

      OIL-PRICE PASSTHROUGH

   ******************************
   THE KEY CONCEPT IN THE PASSTHROUGH ANALYSIS IS THE END-USE
   PRICE OF A COMPOSITE BARREL OF OIL EQUIVALENT.  THE SUM OF
   THE PROCEEDS GOING TO OPEC (OSNOP), TO NON-OPEC OIL (NOSNOP),
   AND TO NON-OIL ENERGY, INCLUDING PROCESSING COSTS, TAXES, ETC
   (SNONOILP), EQUALS THE PRICE.  (IN THE DATA, THIS PRICE IS
   INITIALLY NORMALIZED TO EQUAL 1.  AS THE OPEC PRICE CHANGES,
   THE END-USE PRICE TENDS TO DIFFER FROM 1.)
   :
31:        OSNOP = OFF*RA1+ON*OSNOP(-1)*OPXFPID/OPXFPID(-1)

   IN EQUATION 31, WHEN ON = 1, OPEC PROCEEDS RISE AT THE SAME
```

PERCENTAGE RATE AS THE OPEC PRICE. IN EQUATION 32, A PERCENT INCREASE IN THE SUM OF OPEC PROCEEDS AND PROCEEDS TO NON-OPEC OIL IS CALCULATED. DELTA, AN EXOGENOUS POLICY LEVER, IS NORMALLY EQUAL TO 1. NOSNOP IS NORMALLY HELD EXOGENOUSLY CONSTANT.:

```
32:     PCOPP = (DELTA*NOSNOP+OSNOP)/(DELTA(-1)*NOSNOP(-1)+OSN
        -1))
```

PROCEEDS TO NON-OIL ENERGY ARE MADE TO RISE AT THE AVERAGE RATE OF INCREASE IN PROCEEDS TO OPEC OIL AND NON-OPEC OIL.:

```
33:     SNONOILP = SNONOILP(-1)*PCOPP
```

TOTVAL IS THE RATIO OF THE CURRENT END-USE ENERGY PRICE OF THE COMPOSITE BARREL OF OIL EQUIVALENT TO ITS PRICE IN THE INITIAL PERIOD. IT IS THE SUM OF SIMILAR RATIOS OF THE PROCEEDS GOING TO THE THREE ENERGY SOURCES. BETA IS ANOTHER POLICY LEVER, ALSO NORMALLY EQUAL TO 1.:

```
34:     TOTVAL = DELTA*NOSNOP*OSNOP*BETA*SNONOILP
```

IN EQUATION 35, WHEN ON IS EQUAL TO 1, THE PASSTHROUGH RATE IS THE RATIO OF THE PERCENTAGE INCREASE IN THE END-USE PRICE TO THE PERCENTAGE INCREASE IN THE OPEC PRICE.:

```
35:     PTRS = OFF*RA+ON*((TOTVAL/TOTVAL(-1)-1)/(OPXFPID/OPXFP
        -1)-1))
```

EQUATION 36 CONSTRAINS THE PASS-THROUGH RATE FROM DECREASING OVER TIME.:

```
36:     PTR = IF PTRS LT PTR(-1) THEN PTR(-1) ELSE PTRS
```

CREATIVE PROBLEM SOLVING TECHNIQUES

OVERVIEW

Creative problem solving (CPS) techniques are designed to attack ill-defined problems. That is, they are meant to be used in problem situations in which an obvious way to solve the problem (or even to define the problem) is not clear.

Unlike well-established problem solving methods (e.g., linear programming and network analysis) CPS techniques do not insure that a solution will be found. Furthermore, if a solution is found, CPS techniques do not ensure that it will be the optimum solution. Rather, they are designed help improve chances for finding a solution to problems that, based on standards methods, seem unsolvable.

Most CPS methods work on breaking mind-sets. That is, they help problem solvers to view ill-defined problems from new perspectives. They try to uncover new relationships that tend to be overlooked or discounted from traditional approaches, but which, for the specific problem at hand may have considerable importance. Though differing CPS methodologies utilize differing mental exercises to stimulate idea formation, all have the basic goal of causing two previously well known, but unrelated ideas to be paired in a new or unusual relationship that, in the context of the problem at hand, has useful properties or meaning.

A considerable body of literature exists on the creative process involved in creative problem solving activities, the most well-known being represented by the work of Wallas.[41] According to Wallas, the creative process consists of four stages: (1) preparation, (2) incubation, (3) illumination, and (4) verification. Preparation involves gathering all available information about a problem and would correspond to the intelligence stage in Simon's model. Incubation refers to the time during which conscious effort is suspended but attention is still given to the problem on a subconscious level. Illumination is the "aha" experience that often occurs when a feasible problem solution is recognized. Verification involves testing the validity of an idea and making any necessary refinements.

The major assumptions underlying the formal methods of CPS are specific to their intended purposes. Most techniques for analyzing and redefining problems, for example, assume that these functions can be best performed by factoring problems into their basic elements, examining the interrelatedness of these elements, and achieving perspectives that are remote from the original definition. Most idea-generation methods, in contrast, are patterned into forced-relationship or free-association

---

41 Wallis, G. The Art of Thought. New York: Harcourt, 1926.

processes and assume that deferring judgment and producing a large quantity of ideas are essential functions. However, many of the principles underlying one category of techniques often will apply to another.

CREATIVE PROBLEM SOLVING METHOD: SUGGESTED ADAPTATIONS

I. THE MESS (Identify and Record)

1. Assumptions

2. Biases

3. Conflicts/Concerns

4. Thoughts/Notions (re "the problem" or solutions to "the problem")

5. Goals/Challenges/Objectives/Opportunities (i.e. when I'm finished with this exercise/problem what to achieve--what is my task?)

II. FACT FINDING (Guideline for asking Who, What, When, Where, Why, How)

1. List Headline statements of:

   a. Knowns--facts and opinions

   b. Unknowns--gaps in information

2. In conjunction with a look at "THE MESS", evaluate the FACT FINDING PHASE by placing a check or circle next to the most appropriate items (facts recorded in step #1).

III. PROBLEM FINDING

1. Break "The Problem" down into as many sub-problems as possible-- to develope the broadest possible statement(s) of the problem (Use "I WISH, HOW TO, or HOW MIGHT WE---")

2. Circle or check the most pertinent/appropriate (those which encapsulate the problem) statements to work on

IV. IDEA FUNDING

1. Generate ideas related to "Statements" identified in PROBLEM FINDING. Use Osborn's "Checklist Principles" to "MAGNIFY", "MODIFY", "ADAPT", etc. ideas related to "Problem Finding Statements".

2. Check or circle potential solutions.

SYNECTICS

> This section has been adapted from a portion of the HANDBOOK OF FORECASTING TECHNIQUES, Part II, Description of 31 Techniques, prepared for the U.S. Army Engineer Institute for Water Resources, Fort Belvoir, VA. The work was done under Contract No. DACW31-75-C-0077, August 1977.

ABSTRACT

Synectics is a well-tested method for developing creative solutions to problems in business, government, technology, and people situations. It could probably be used in conjunction with other techniques to generate alternative futures, forecast trends, and enrich cross-impact analyses. Courses are offered in the technique which enable organizations to explore the applicability of Synectics to classes of issues of concern to the organizations.

DEFINITION

Synectics is described by its developers as a technique for "dynamic group problem solving." The name is a Greek word meaning the joining together of different and apparently unrelated elements into a synergistic whole. The total process is a means of applying structured creativity to a huge array of technical, business, governmental, and people problems. In a Synectics session a specific problem is defined. The group (usually less than ten) includes a trained Synectics leader, an expert in the problem area under study, and random others who hopefully possess reasonably unfettered minds. Under the guidance of the leader, the group develops and discusses analogies to the problem. Having done so in depth, the leader restates the initial problem and a "force fit" is achieved between the problem and the analogy. The basic purpose is to force people into fresh viewpoints based upon (often far fetched) analogies to the problem calling for solution and in this fashion apply information stored in corners of the conscious and subconscious mind to the issue at hand. (Further details appear later under "Procedures.")

PROBLEM SOLVING TECHNIQUES

Creative Problem Solving

HISTORY

Synectics as a problem-solving technique was developed in the 1950s by W. J. J. Gordon. In 1961 he presented his basic approach in a book called Synectics. A company, Synectics, Inc., was founded in 1960 in Cambridge, Massachusetts, to exploit the approach. Since then a great diversity of problems calling for creating solutions have been tackled. Much of this work--including fascinating transcripts of many actual sessions, some successful and others not--is described in The Practice of Creativity by George M. Prince (Harper & Row, 1970). Mr. Prince is president of Synectics, Inc.

MAIN USES

Synectics has been applied in areas ranging from new-product hardware problems, to business process problems, to people-oriented process problems. The list below cites specific examples in these areas:

New-Product Problems--Hardware-Oriented

Devise a more efficient fuel cell.

Invent an easier way of applying paint.

Invent a better means of closing a Thermos bottle.

Devise an instant, portable radio antenna thirty feet tall that travels in a small package.

Devise a profitable use for a waste by-product.

Expand an interesting piece of technology into a commercial product line.

Conceive of new home appliance that fulfills a need that no one is now aware of.

Discover how to measure the oil content per unit of stone in a deep formation.

In one model of an army personnel carrier, land mines tend to trigger a gasoline explosion in the fuel tanks. How can this be prevented without major alteration?

Process Problems--Business-Oriented

Devise a continuing education program in a company--one that will keep employers interested and alert and will avoid obsolescence.

ANALYSIS ACTIVITIES

Creative Problem Solving

Conceive of a more effective method for acidizing oil-bearing limestone strata.

How can we inexpensively add one pound of chemical to two tons of grain and have each grain get its fair share?

Devise a new market strategy for a dying brand.

How can a manager in a technical area be most helpful when a subordinate has a technical problem?

Devise a system for presenting ideas that gives then a maximum possibility for constructive consideration.

Devise an idea-incentive system that will encourage involvement on the part of everyone from janitor to vice-president.

Process Problems--People-Oriented

How can a bored clergyman renew himself?

How can we make the visiting-physician program in Vietnam produce more lasting benefits for Vietnamese physicians?

How can an individual reduce his prejudices?

Conceive of the physical facilities of a new architectural school that satisfies the needs of the interested parties (students, faculty, towns-people, etc.)

Devise an economical system in which both a slow and a fast student can be given what each needs in the same class.

How can we persuade those in power in the ---- to pass this power downward in a orderly fashion?

Conceive a meaningful way to involve people in the democratic process from childhood on.

On a personal level, people familiar with the techniques of Synectics often claim to be better able to deal with everyday problems of living, as well as with work-oriented issues.

LIMITS AND CAUTIONS

As in any creativity endeavor, results are unpredictable. In Synectics, the "forced fit" between analogy and problem (e.g., between the way a cat acts and a deep oil-bearing stratum) results in what is called a "viewpoint," which sometimes points to a spectacular solution to the problem and is sometimes a dud. Proponents of the system like to refer to

Synectics as "dependable creativity," but this is certainly putting it rather strongly.

In theory Synectics can be applied to any area of interest, provided there is such a thing as an "expert" to state the problem and evaluate the feasibility of viewpoints arising from "forced fits." For purposes of forecasting alternative futures, conducting cross-impact analyses, or for projecting trends in specific areas, it seems possible that some kind of amalgamation of Delphi and Synectics principles would help overcome shortcomings of both approaches.

OTHER TECHNIQUES

Many other techniques strive for creativity. These include brainstorming, case methods, hypothetical situations, morphological analysis, value engineering or value analysis, attribute listing, bionics, forced thinking in specified channels, life way analysis, "genius" scenarios, utopian or "what if" thinking, and many more.

PROCEDURES

Synectics sessions can be as brief as an hour or several days in length. As indicated earlier, the group is generally limited in size and requires the presence of a leader trained in Synecticts and of a subject matter expert. As Gordon, the principal developer of the approach, says "The synectics leader is primarily responsible for keeping the investigation of the problem within the confines of the synectics flow chart and ensuring the most efficient generation, development, and use of analogical material. Which analogical route to take is an important decision. It is made by the leader on the basis of the criterion of constructive psychological strain. In a people-oriented problem, this means that the leader would seek analogies from the exact sciences. In a mechanical problem, he might look to biological models."

The steps in the Synectics process are:

1. Problem as Given
   Brief explanation of the problem as given. A general statement is made of the problem to be solved as it may have been given to the group members by an outside source or as generated by themselves. Some examples:

   Invent a wheelchair that will go up stairs.

   Design a nonfogging bathroom mirror.

ANALYSIS ACTIVITIES

Creative Problem Solving

Build an anchor that will have more holding power per pound of weight than anything now available.

Develop a charcoal that will ignite faster.

Compress food.

Close openings more effectively.

Discover new stimulants for the healing process.

2. Analysis and Evaluation by Expert.
An explanation off the problem is presented by the most expert in the group, making the strange problem familiar. Enough detail is given about the problem to permit the group to go to work. The expert is a participant. He does not need to try to make everyone as knowledgeable as he is.

3. Purge.
The most usual solutions are voiced, evaluated, and purged. Getting these immediate solutions out for an airing permits progress toward more unique and more valuable solutions.

4. Generation of problems as understood.
After the given problem has been explored, each participant writes a restatement of the problem as he sees it or a goal he believes would be desirable. It is considered useful to write several versions of the problem which imply different approaches to it.

5. Choice of problem as understood.
A momentary agreement is reached by the group on an understanding of the problem.

6. Use of evocative questions.
The leader employs evocative questions to stimulate the employment of the analogical operational mechanisms.

7. Examination.
A possible solution stemming from stage six is selected by the group for examination. The cycle continues, using evocative questions to stimulate the use of imagination to come up with other solutions and then the examination of these solutions.

8. Force-fit.
The solutions produced by the foregoing process are forced into fitting the nature of the problem originally posed. The attempt to fit the tentative solutions to the problem may result in the problem itself being seen differently and may suggest new lines of speculation.

9. Viewpoint.
From the material arising in the force-fit stage, a new viewpoint of the problem is reached, and many potential solutions may arise. One

of the basic differences between the Synectics method of operation and traditional problem solving methods is that the latter seek solutions more directly. Synectics seeks new lines of speculation, and these, in turn, lead to potential solutions by means of the force-fit.

PRODUCT OR RESULT

The upshot of the Synectics process is a new way of looking at the initial problem. These are as diverse as ideas and can be presented in as many ways.

LEVEL OF DETAIL OF RESULTS

Synectics produces a viewpoint--an approach--rather than an "operative answer." In most cases operational details remain to be worked out, although they may be almost self-evident. For example, Synectics hit upon the idea of compressing Kleenex to reduce shipping costs, of freezing cores prior to lifting to study in situ oil conditions, of coating tire rims with epoxies to prevent leakers (a problem because slow leaks caused closely parked cars to fall against each other and be damaged in transit). In each case many technical details had to be worked out following the initial insight arrived at through Synectics.

LEVEL OF CONFIDENCE OF RESULTS

In advance of a Synectics meeting, one can only hope for "break-through" results. Viewpoints resulting from a meeting range from self-evidently workable to 99% certainty that the solution won't work.

COMMUNICABILITY OF RESULTS

Usually excellent because the central results of Synectics tend to be non-technical universals understood by everyone. This, of course, does not apply to descriptions of the nature of the problem (often highly technical) or to implementation of Synectics-based solutions (again often technical).

CREDIBILITY OF RESULTS TO CRITICS

Many people trained in the exact disciplines are uncomfortabl: with creativity techniques and hence tend to downgrade Synectics. These same critics, however, would certainly admit that Synectics has produced some highly useful results, although many might argue the result could have been arrived at (perhaps faster and with greater assurance of success) through conventional channels.

SPAN OF FORECASTS

As applied to social or technological forecasting (perhaps in collaboration with Delphi methods), the forecast span would tend to be mid-range or longer.

RESOURCES

Resources are wholly in the form of people, although a recorder is usually used to tape sessions.

Key personnel are the Synectics leader, the subject expert, and others selected for high energy levels, broad job and/or educational backgrounds, and openness of thinking mode.

Sessions devoted to a major problem typically take two or three days from start to finish. Costs are chiefly for time.

ASSESSING AND PREDICTING DECISIONS

INTRODUCTION

> This section provides an overview of decision processes, both intuitive and methodological. It is intended to give you, as analysts, a general feel for the kinds of factors that might affect the decision process (and hence the decisions) of foreign decision makers or decision making groups. It is not intended to make managers or decision makers out of you (OTE has management courses designed for that purpose). Use this lecture to gain a general understanding, not skills level capabilities.

METHODS AND MODELS FOR ASSESSING AND PREDICTING DECISIONS

Making a decision does not equate to solving a problem. The decision also needs to be implemented for it to have a chance of solving (or at least reducing in intensity) the problem being attacked.

Implementation methods can vary dramatically from one problem area (decision maker) to another. Heads of state usually have the authority and power to see their decisions implemented. Other officials often must seek the approval of others, or try to gain the needed support by influencing, coercing, or begging. To understand or predict the likely outcome of a pending decision, we must, therefore, know something about the players involved and the power dynamics of the situation. We also will be better off if we know something about the decision style (predictable or innovative, for example) of each of the players or decision groups.

In formulating decision problems, a decision maker(s), chooses among available strategies or options that help him (her) compete against either an uncertain future or an opponent who will anticipate his actions and react. The decision maker's selection of a strategy coupled with the future occurrence of a specific state-of-nature or an opponent's selection of a strategy determine the outcome of the decision. This outcome will have some specific "goodness" or "badness" associated with it if measured in terms of the decision maker's objective(s).

ANALYSIS ACTIVITIES

Assessing and Predicting Decisions

There are several viewpoints (or models) available for examining decision processes. Three useful ones[42] are:

1. The Rational Actor -- where the decision maker(s) is viewed as an individual or coherent (single purpose) group that acts in a manner that is consistent, in terms of the decision maker's goal(s), with maximizing the expected return (or minimizing the expected loss) that will result from a decision.

2. The Political Actor -- where the decision process is assumed to be shared by multiple players having competing (often conflicting) decision objectives. Each player strives to maximize his own personal benefits from the "group" decision; with the result that the final decision may not be optimal (rational) in terms of the objectives of the umbrella organization common to all of the decision players.

3. The Bureaucratic Actor -- where the objectives of the decision maker are defined more strongly by his (her) position in the organization and the rules governing acceptable and expected behavior within the organization, than by the objectives that would define a optimal decision to the organization itself. (A special case of this model describes a decision maker working in the public sector, where service and accountability are frequently the criteria for evaluating performance, rather than profit, as is more frequently the case in the private sector of an economy.)

The following two sections provide methodological structures that can incorporate the rational, political, and bureaucratic models of decision making into mathematical frameworks developed around:

(1) the manager's choices

(2) the possible future states-of-nature or opponent's choices

(3) the worth (to the decision maker) of the possible joint occurrences described in (1) and (2).

The first structure to be discussed is that of Decision Theory or Decision Analysis, which generally deals with cases where the manager selects strategies which are affected by states-of-nature. The assumption is that what the manager does has no impact on which state-of-nature will occur, but that the state-of-nature can have significant impact on the goodness or badness of the outcome of the manager's choice. Thus, the state-of-nature is considered to be an exogenous variable; it affects the system, but the system does not affect it. (Note that the future state-of-nature may be single valued; that is, it may be certain. Decision analytic techniques still apply under conditions of certainty, they simply "degenerate" into straight-forward multiattribute utility assessments. When the

---

[42] These three models are developed in terms of the Cuban missile crisis by Alison in <u>The Essence of Decision</u>.

future is uncertain, these same multiattribute utility assessment techniques are combined with probability theory to determined an expected value for each decision alternative that appropriately weights each possible future outcome. This is decision making under conditions of uncertainty or risk.)

The second structure discussed is that of the Theory of Games, or Game Theory. This set of methodologies generally analyzes situations where the decision maker and an opponent(s) both select strategies in competition with each other. The possible combinations of strategies chosen simultaneously (though independently) by all players determine the value of the outcome to each player. The assumption is that there is interaction between the decision maker and an opponent with regard to their selection.

Notice that by assuming that either there is an interaction between the decision maker and an opponent or that the decision is made against an uninformed and disinterested "mother nature" distinguishes between these two general sets of methodologies: Decision Analysis and Game Theory. As both of these methods are covered, hopefully you will seen that the two methods are in fact intertwined and useable as a collective set.

ANALYSIS ACTIVITIES
Assessing and Predicting Decisions

DECISION ANALYTIC TECHNIQUES

OVERVIEW

Decision analysis is basically a hierarchical structuring procedure that uses utility theory to assign values to outcome events or acquisition choices. Probability theory, making heavy use of subjective probability assessments, is used in conjunction with the assigned utils (regrets) to obtain expected values for decision choices made under conditions of uncertainty. Under conditions of certainty (e.g., buying a new car, selecting the best contractor proposal, or -- the examples to follow -- evaluating employee performance or allocating resources most effectively) the utils are assigned to multiple attributes of the objects or decision outcomes being chosen among; and the expected value (regret) of each decision option is determined by combining the utils assigned to each of the attributes via a system of weighting factors similar to the one discussed earlier in this text in the Information Structuring section (page 75). Similar weighting schemes are used to evaluate the relative merit of competing projects that could be funded from a budget too small to fund all possible projects.

Decision analytic techniques are frequently time consuming to exploit fully. But, they provide extremely through documentation for how a given decision was reached, and also provide a great deal of insight (to the decision maker) of why the decision was made and of what the relative trade-off is among alternative decision choices.

UTILITY THEORY

Utility theory deals with developing numerical scales that express the psychological value of objects or outcomes to someone, usually a decision maker. For instance, developing a numerical scale that expresses in terms of interval scale numbers the relative goodness or badness of possible treaty alternatives. Often utility scaling involves the "rescaling" of what at first glance seem to be interval scales already. For instance, people usually find that the unit value of a dollar depends on whether the dollar is the only one in question, or one out of 10, or out of 1,000,000. That is, the difference between zero dollars and one dollar often is evaluated to be a more important difference than the one dollar difference between $1,000,000 and $1,000,001. Similarly, obtaining a fifth Rolls Royce probably would mean less to you than obtaining your very first one.

The trick to expressing these psychological values in numerical terms is the fundamental concern of utility theory, and many differing approaches, both for obtaining the utility (or regret) values and for using them, have been developed. We'll discuss a brief sample of these techniques, mostly for the purpose of familiarizing you with the meaning of util and regret scores, so that the decision analysis discussion to follow will be more interpretable.

EXPECTED VALUE

Expected values are a key concept in probability theory, statistics, and game theory. The expected value of a strategy is obtained by estimating the probability of each possible outcome, and multiplying each of these probabilities times the value (utility) associated with that outcome. The sum of these multiplications is the expected value of the strategy, and a risk-adverse player will normally select the course of action with the highest expected value. For example: Assume that Strategy I has the following three possible outcomes (with the associated probabilities of happening):

Outcome	Probability	Payoff
A	0.6	30
B	0.2	80
C	0.2	-20

Also assume that Strategy II has two possible outcomes with the indicated values and probabilities of occurrence:

D	0.5	60
E	0.5	-10

The expected value of Strategy I is:

$$(0.6 \times 30) + (0.2 \times 80) + (0.2 \times -20) = 30$$

The expected value of Strategy II is:

$$(0.5 \times 60) + (0.5 \times -10) = 25$$

ASSESSING RELATIVE VALUE UNDER CONDITIONS OF CERTAINTY

ASSESSING THE RELATIVE WORTH OF MULTIPLE THINGS JUDGED CN MULTIPLE CRITERIA

Summary
This section outlines a personnel evaluation procedure based on decision analytic, relative-weighting multi-attribute utility assessment techniques. An example assessment of one group of employees is included to illustrate how the procedure might be implemented and how information about intermediate results could be presented to to facilitate the rating process.

ANALYSIS ACTIVITIES

Assessing and Predicting Decisions

Introduction

The merit or worth of people or things can be assessed in many ways. In everyday matters unaided intuitive judgment usually is appropriate. At times, though, a more structured approach may be desirable, either because the assessments are complex or because an explicit audit trail of how the final conclusions were reached is desired.

Hierarchical assessments add both structure and documentation. They involve indicating, in an explicit manner, both the identity and the relative importance of the factors that contribute to the assessment process. For instance, an employee's value might be disaggregated into criteria such as initiative, reliability, creativity, and judgment. This disaggregation process can be continued for multiple levels by further disaggregating each of these criteria into more basic components (e.g., Creativity might be disaggregated into "Idea Generation" and "Information Synthesis").

Disaggregation normally is stopped when assessments of the "bottom nodes" can be made directly. For the purposes of personnel evaluations, a single level of disaggregation is recommended. It should provide sufficient detail to document the evaluation process, without requiring excessive time investments.

Two things must be done to produce an overall assessment of personnel from any hierarchical structure developed.

1. Each person being evaluated must be scored in each criterion used in the assessment structure. A bottom-up, relative weighting scheme is used. This procedure (to be discussed in detail later) involves comparing the persons being assessed against each other directly (within each assessment criterion separately), rather than against an absolute scale that must be developed separately.
2. The relative weight that the scores from a given criterion will contribute to the overall assessment needs to be established. The bottom-up procedure again uses direct comparisons (of the range in quality within each criterion), rather than a global assessment of the generic value of a criterion assessed in the abstract.

This procedure emphasizes the factors that best discriminate among the employee characteristics being evaluated, and makes explicit assessments of the relative strengths of each employee (on a criterion by criterion basis) compared to the other personnel rated.

Developing an Evaluation Structure

The first step to developing an evaluation structure is for the reviewing panel to identify the job types (e.g., project officer, supervisor, secretary) and other categories (e.g., management potential, upward mobility potential) to be used in evaluating personnel. For each evaluation category the specific set of criteria to be used for evaluations in that category must be identified.

This selection of a specific set of (probably 10-15) criteria, tailored specifically for each evaluation category, is based on the assumption that some criteria will be useful only for evaluating secretaries (e.g., typing skills, shorthand), whereas other criteria will apply only to supervisors (e.g., counciling skills, administrative knowledge). Some criteria (e.g., initiative and judgment) may be relevant for all evaluation categories. Table &tabl provides an initial "shopping list" of criteria potentially useful for GS-15 evaluations.

TABLE 19 -- SHOPPING LIST OF EVALUATION CRITERIA FOR GS-15'S

Initiative	Reliability
Enthusiasm	Consistency
Self-Awareness	Flexibility
Responsiveness to Direction	Desire to Improve
Creativity	Technical Knowledge Depth
Preparation	Technical Knowledge Breath
Planning	Writing Skills
Priority Setting	Speaking Skills
Perspective (Big Picture)	Ethics
Judgment	Methodological Management Skills
Timeliness	Administrative Knowledge
Efficiency	Leadership Skills
Productivity	Counciling Skills
Throughness	Technical Growth Potential
Tenacity	Managerial Growth Potential
Compatability (Impact on Work of Others)	

This list must be upgraded prior to the selection of criteria. Primarily, the criteria must be defined clearly. This will identify criteria that have multiple dimensions (can be interpreted in more than one way), so that such criteria either can be disaggregated into uni-dimensional terms, or defined such that only the desired dimension will be used for assessments.

The specific set of criteria chosen for use in any given evaluation category should be selected on the basis of two factors:

1) the generic importance of each criterion to the attributes associated with the evaluation category, and
2) the degree to which the criterion can be used to distinguish among the employees being rated in that category.

That is, if a criterion (initiative, for example) is considered to be important to assessing project officer performance, then it passes the first test. But, if the rating panel judges that all project officers have exactly the same level of initiative, then it would fail the second test, since it would provide no evaluative discrimination among the project officers being rated.

An Illustrative Example

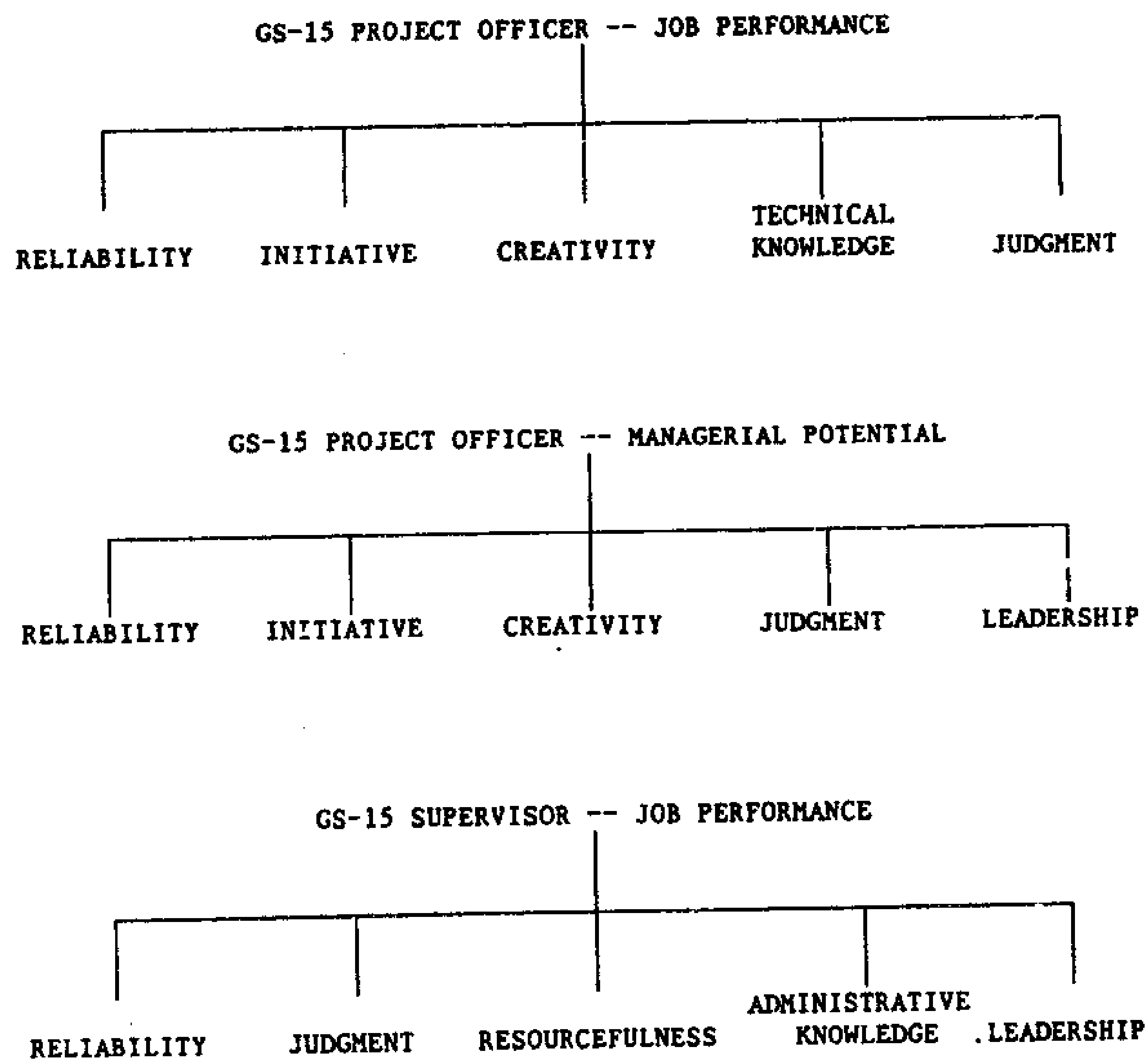
The Rating Structure.
Assume, for purposes of an example, that a Career Service Panel (CSP) decided that five criteria were sufficient to evaluate personnel, and that the three criteria sets shown in Figure 35 were chosen to evaluate GS-15 supervisors (on job performance) and project officers (separately on job performance and management potential).

Note first that there is overlap among the criteria sets. All of the evaluation categories include Reliability and Judgment; but only the supervisory category contains Resourcefulness and Administrative Knowledge. Similarly, only the project officer category includes Technical Knowledge, and the supervisor and management potential categories both contain leadership.

The Rating Forms.
If we assume that only project officers need to be evaluated for management potential (since supervisors are already being rated on their performance as managers) then the project officers need to be scored against each other on the basis of six criteria, and the supervisors on the basis of five. The six criteria used for the project officer evaluations will be subdivided into two (overlapping) groups of five to assess separately a) job performance and b) management potential; with the relative weights given to each criterion probably differing between the two evaluation categories (the procedure for doing this will be discussed below). Sample versions of forms that a CSP member might use (independently) prior to a CSP evaluation meeting to assess the hypothetical GS-15 population of this example (there are seven project officers and 5 supervisors) are shown in Figures 36 and 37.

Within Criteria Scoring.

FIGURE 35 -- GS-15 EVALUATION CRITERIA



To use the sample forms, each CSP member first would rate each person against the others in their group <u>within each criterion separately</u>. For a bottom-up scoring process, the rater first identifies the best and the worst person from the group in each criterion category. The worst person is then assigned a zero in that criterion, and the best person a 100. The remaining people are assigned scores between 0 and 100 indicating where they rate (in terms of the criterion being assessed) relative to the best

FIGURE 36 -- SAMPLE FORM FOR CSP EVALUATIONS OF PROJECT OFFICERS

GS-15 PROJECT OFFICERS

	RELIABILITY	INITIATIVE	CREATIVITY	TECHNICAL KNOWLEDGE	JUDGMENT	LEADERSHIP
Able						
Baker						
Charlie						
Dave						
Earl						
Frank						
George						

CRITERION WEIGHTS:

Job Performance						-0-
Management Potential				-0-		

and worst persons. This procedure is repeated separately and independently for each criteria.

For example, assume that in rating the project officers on initiative, it was felt (by the CSP member doing the rating) that Earl was the

FIGURE 37 -- SAMPLE FORM FOR CSP EVALUATIONS OF SUPERVISORS

GS-15 SUPERVISORS

	RELIABILITY	JUDGMENT	RESOURCEFULNESS	ADMINISTRATIVE KNOWLEDGE	LEADERSHIP
Herman					
Igor					
Jed					
Ken					
Lyman					
CRITERION WEIGHTS:					

---

strongest candidate on the basis of initiative, whereas George was the worst. Earl would therefore be given a a score of 100 in the initiative column, and George would get a 0. The other five people on the project officer list would now be assigned scores intermediate between 0 and 100 to indicate the CSP rater's assessment of where each person's level of initiative fell compared to Earl's and George's. If, for instance, it was felt that Charlie's initiative placed him about a quarter of the way from George toward Earl, then he would be assigned a score of 25. Able would get a score of 75 if his level of initiative extended three quarters of the difference between George and Earl. These sample scores are shown in Figure 38.

Ties are allowed. More than one person can be given a 0 or a 100 or any number in between. However, at least one 0 and one 100 must be assigned (unless the rater feels that all of the people being rated have exactly the

FIGURE 38 -- ILLUSTRATIVE COMPLETED EVALUATION FORM FOR GS-15 PROJECT OFFICERS

GS-15 PROJECT OFFICERS

	RELIABILITY	INITIATIVE	CREATIVITY	TECHNICAL KNOWLEDGE	JUDGMENT	LEADERSHIP
Able	50	75	80	60	45	30
Baker	30	45	100	55	80	50
Charlie	0	25	15	0	10	0
Dave	100	85	90	100	100	100
Earl	80	100	100	95	70	90
Frank	60	55	50	85	85	60
George	15	0	0	20	0	5

CRITERION WEIGHTS

Job Performance	70	20	25	75	100	-0-
Management Potential	100	40	20	-0-	95	80

same rating in that criterion, in which case the criterion should be dropped--by giving it a zero criterion weighting.

The numbers assigned do not have meaning yet in absolute terms; they indicate relative strengths of each person within a given attribute. That is, the 0 (zero) assigned to George for initiative does not imply that he has no initiative; nor does Earl's 100 imply that he he is the ultimate role model for initiative. Instead, the numbers relate to the degree of satisfaction that the rater would feel if he had to trade off one of the personnel (in this group) for another in terms of initiative. The rater would get no (i.e., zero) satisfaction in terms of initiative if George were assigned to work for him (since George is the worst option, in terms of initiative, that could happen). The rater could do no worse. Conversely, if Earl were made available, the rater would achieve the maximum satisfaction (100%) possible from this group of project officers. The 75 assigned to Able implies that the rater would achieve 75% of the satisfaction (in terms of initiate) if Able were made available compared to the alternative choices of Earl and George.

The fact that 0 to 100 scores are generated in each criterion category does not imply that each criterion is equally important. The relative importance of the criteria must be established separately.

Establishing Criteria Weights.

To give the 0 to 100 range of scores in each criterion its appropriate impact on the overall assessment of the personnel being evaluated, the relative importance of the differences within each criterion category must be established. To do this, the CSP rater must compare the difference between the best and worst person in each criterion category against the corresponding difference in each other criterion. Each criterion is then assigned a weight to indicate the relative significance of the difference in satisfaction between the best and worst person in a given criterion category, to the corresponding difference in each of the other categories. The rater does not simply judge which criterion is generically the most important, but rather judges where the most important difference exists for the specific group of employees being rated. For instance, if the rater felt that the most important attribute for a project officer to have was initiate, this criterion category still could be assigned a relatively low importance weight if all of the project officers in the group being evaluated had about the same degree of initiative.

One of the easier ways of assigning importance weights is to first determine which of the criterion has the biggest difference between the worst and best person. To do this, the CSP rater basically answers the question: Given the choice between going from George to Earl in terms of initiative versus going from "the worst" to "the best" person in another criterion, which improvement would I choose first. Pairwise comparisons are made in this fashion until the top choice of all criteria is determined.[43]

---

[43] These pairwise comparisons must be made separately for each evaluation category to which the personnel being assessed belong, because the trade-off question might be answered differently if assessed from the standpoint of job performance as a project officer (where technical knowledge might rate very high), compared to assessing from the viewpoint of judging management potential (where technical knowledge might have no importance at all).

Let's consider weighting assessments from the standpoint of job performance. If the intracriterion swing in job performance is judged to be greatest in the category of judgment, then the rater would assign some number (100 is usually convenient to use) to the Judgment column for Job Performance. Each of the other criterion categories would then be compared to the Judgment one, and numbers would be assigned to indicate the fraction of importance associated with the difference in that criterion compared to the difference assessed for Judgment. This weighting process differs from the within criterion scoring process discussed above in that criterion category is required to be assigned a zero. In fact, this should occur only rarely (except for those criterion factors which were excluded, see Figure 35, from use in assessing a given evaluation category).

Figure 38 shows illustrative criteria weightings for both the Job Performance and the Managerial Potential evaluation categories. The zeros in each row reflect the exclusion of Leadership and Technical Knowledge as relevant evaluation criteria for Job Performance and for Managerial Potential, respectively. To check the consistency (subjective veridicality) of the weights, the CSP rater compares criterion categories in pairings (or other combinations) not used to assign the weights initially. For example, the 70 and the 75 assigned to Reliability and to Technical Knowledge, respectively, imply that the difference between best and worst within each of these two criteria were assessed as about the same importance, with Technical Knowledge being slightly more significant a difference. Similarly, the numbers also imply that the combined difference assessed to exist in both Creativity and in Technical Knowledge, when considered together, should be about equally important as the difference in Judgment alone. If these consistency checks seem right, then the CSP rater is finished with this individual rating process; otherwise, the weights should be reviewed and revised until the weights are consistent for all possible comparisons.

Once a consistent set of rating scores and and criteria weightings has been received from all CSP members, they can be used to provide summarized feedback and discussion guidance for the CSP group rating process that follows.

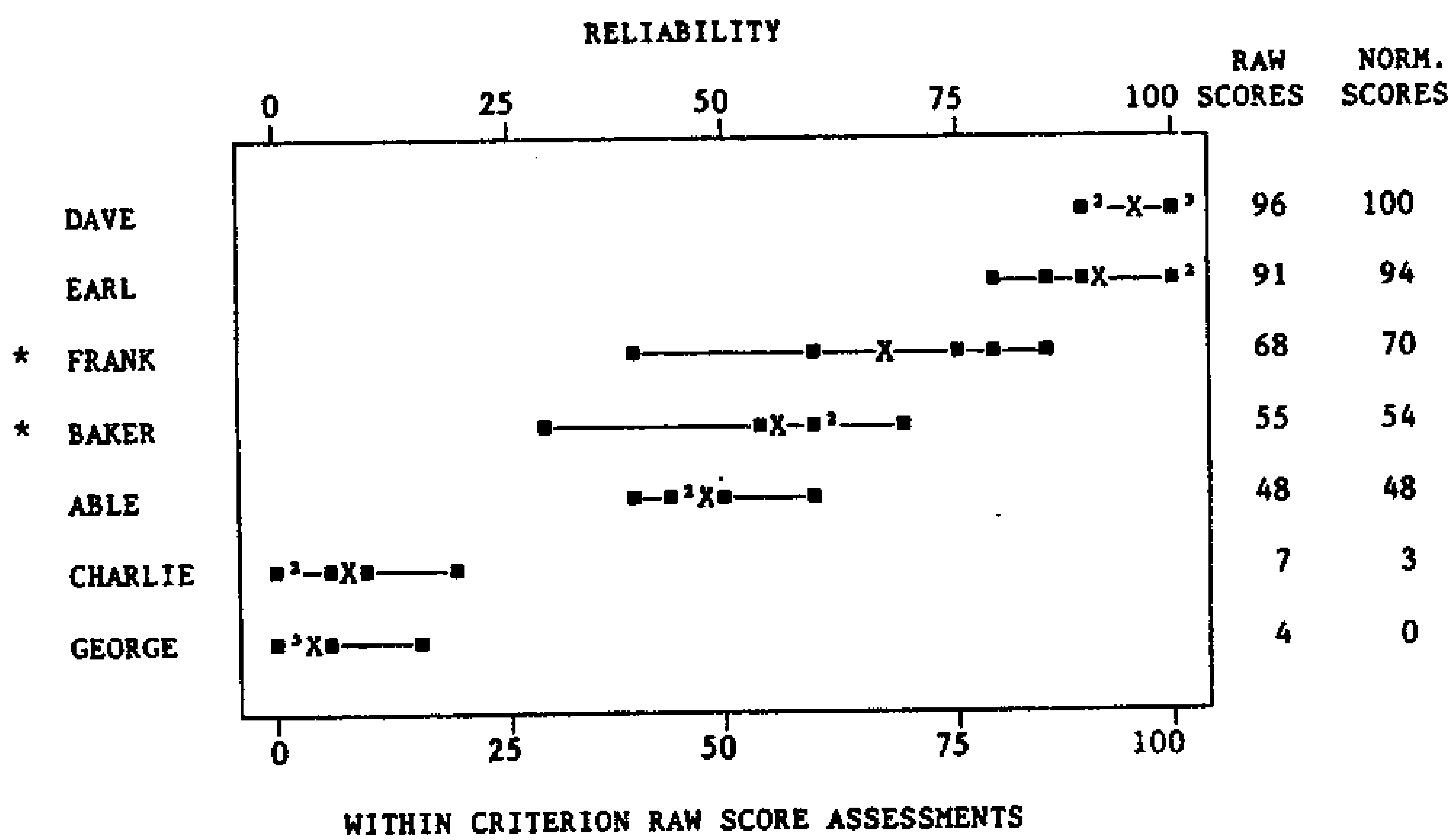
Feedback from the Independent Ratings

The primary purpose for collecting the independent ratings from each CSP member is to provide a starting point for joint discussions to finalize the rankings. The bottom-up weighting process will permit more than just a summarized, rank-ordered list to be provided. It also will enable areas of disagreement (where discussion is needed) to be identified explicitly. By default, areas of agreement, where extended discussion is <u>not</u> needed, will be made clear also.

Both graphical and tabular data presentations can be produced quickly from the raw data of the independent rating scores and criteria weightings. For example, Figure 39 shows one possible display that presents graphically the 5 independently scored ratings given to each GS-15 Project Officer in the Reliability Criterion. The spread of the scores given each analyst indicate the degree of agreement among the CSP members in rating this criterion. The asterisks in the left margin flag those people (Frank and

Baker) where disagreement (defined, for this example, as a difference of more than 25 between the lowest and highest score given to an individual) among CSP raters was present. The Raw Score column at the right gives the algebraic mean of the scores received, and the Norm(alized) Scores column gives the scores that result by converting the Raw Score column numbers into a 0 to 100 scale. A graphical aid such as Figure 39 should help the CSP to collectively assess the individuals in this criterion and produce a concensus set of Reliability rating scores to use for final evaluation purposes.

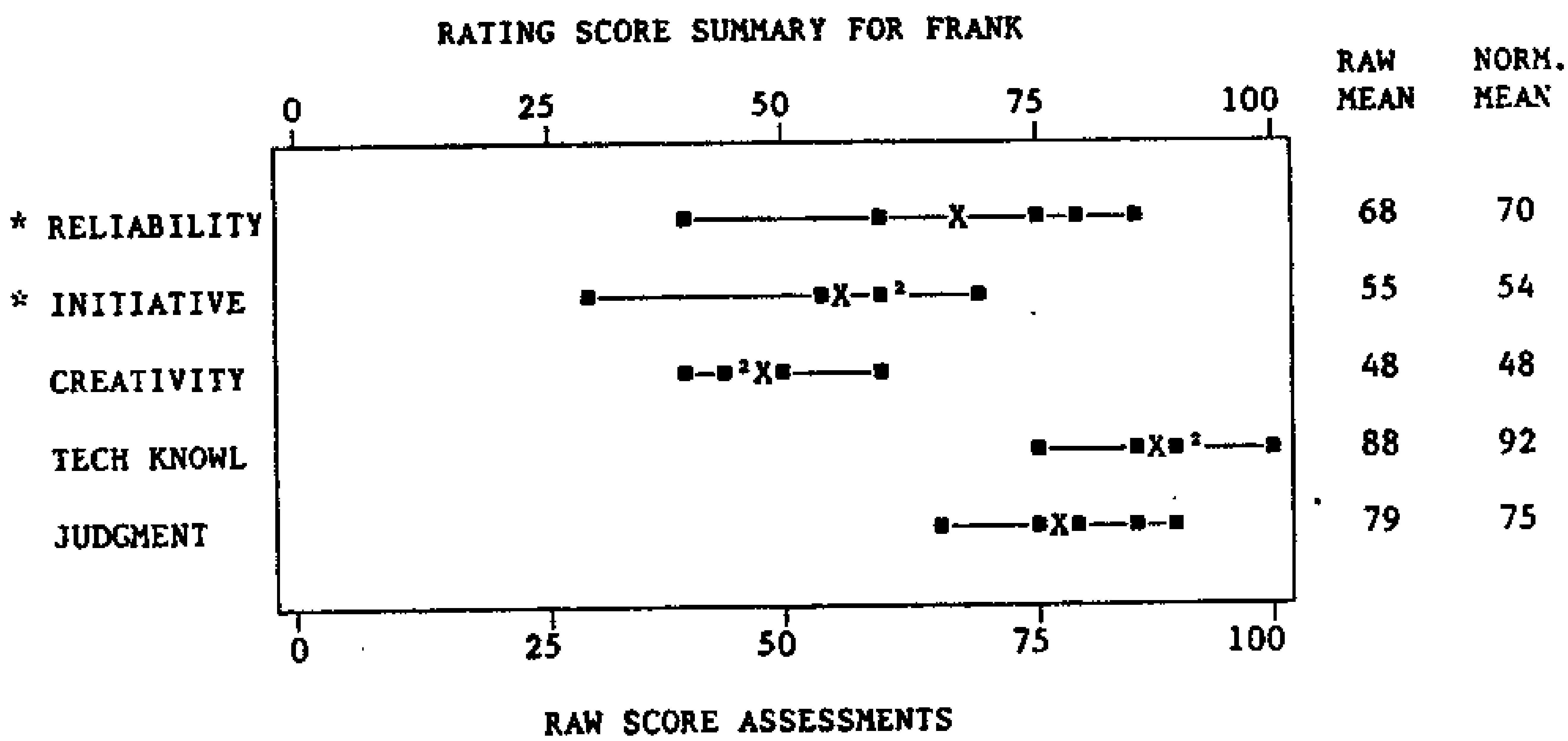
FIGURE 39 -- GRAPHICAL SUMMARY OF RELIABILITY SCORES FOR GS-15 PROJECT OFFICERS



The superscripted numbers indicate multiple scores assessed at that point. The asterisks at the left indicate greater than 25 point disagreement between the highest and the lowest scores assessed that individual. The "X"'s indicate the algebraic mean of the individual scores received. The "■" show the individual scores.

Another display (partially redundant with the first) is shown in Figure 40. It shows graphically a summary of the independent scoring results for each person; thus indicating an individual's consistency, as well as general level of performance, across all rating criteria. Asterisks are used again to flag significant disagreement among the CSP raters. Since poor performance in any given criterion is clearly evident from this type of display, training alternatives for "teachable" criteria are identified also.

FIGURE 40 -- GRAPHICAL SUMMARY OF SCORES RECEIVED BY ONE GS-15 PROJECT OFFICER



Superscript numbers indicate multiple scores assessed at that point. The asterisks at the left indicate greater than 25 point disagreement between the highest and the lowest scores assessed for that criterion. The X's indicate the algebraic mean of the individual scores received. The ▪ show the individual scores.

The display illustrated in Figure 40 would not be as directly useful for CSP group rating purposes, since it does not provide the comparative information among individuals that is is needed for the within-criterion scoring. It does, however, give CSP members some initial information about which personnel are perceived as doing poorly or are rated inconsistently by CSP members. This should give the members a chance to prepare or collect information that they want to bring to the attention of the CSP about persons felt to have been assessed inaccurately. A Figure 40 type of display also could be useful to display how each individual did in the finalized CSP ratings.

Figure 33 shows how the overall results of the initial, independent ratings could be formated in a table similar in appearance to the rating forms used, but listing people in order of the overall ranking calculated from the numerical input to the hierarchical rating structure.**

Graphical displays also can be used to facilitate CSP deliberations to establish concensus criteria weights. Figure 42 shows how the criteria weights assessed independently by the CSP members can be normalized and displayed graphically to flag categories in which there is significant disagreement. After group CSP deliberations have produced concensus within criterion scores for all criterion and all people within any evaluation category, a graphic such as Figure 42 should help to focus discussions for reaching concensus criteria weightings. (The within criterion scores must be deliberated first, so that the concensus best and worst individual in each criterion are known prior to comparing the best-worst difference from one criterion to that of another.)

The CSP Group Refinement Process

Several methods could be used for the CSP to refine the initial, independent ratings. Based on the feedback from the initial data, members could revise and compare individual rankings iteratively (in a DELPHI-like process) until major disagreements were resolved. Alternatively, members could simply use the initial feedback for reference and, via a group decision process, produce a single set of concensus criterion scores and weights. The latter technique has the advantage that sensitivity analyses of the impact of changing assessments can be conducted relatively easily. That is, the impact on the order and spacing of the overall individual rankings that results from changing a criterion weight by a factor of 2 or 3 can be readily calculated and displayed. This can be quite useful in cases where the sensitivity analysis shows that the range in some weighting factor about which the CSP is having difficulty reaching concensus has no

---

** The calculations for a single level hierarchical structure are simply an algebraic, weighted average of the concensus criterion scores. The normalized, concensus criteria weights are simply multiplied times the concensus score assigned an individual in that criterion, and this product is added to the comparable (score times weighting factor) products from the other criteria to determine the overall score. These overall scores will range between 0 and 100, inclusive; with a 0 or a 100 overall score being possible only if an individual is assigned a 0 or 100, respectively, in each and every criterion category. If the individuals are being evaluated in multiple categories (e.g., Project Officer and Management Potential) then each person will receive an overall ranking, based on the appropriate set of criteria weights, in

FIGURE 41 -- TABULAR SUMMARY OF CSP EVALUATIONS

GS-15 PROJECT OFFICERS -- JOB PERFORMANCE

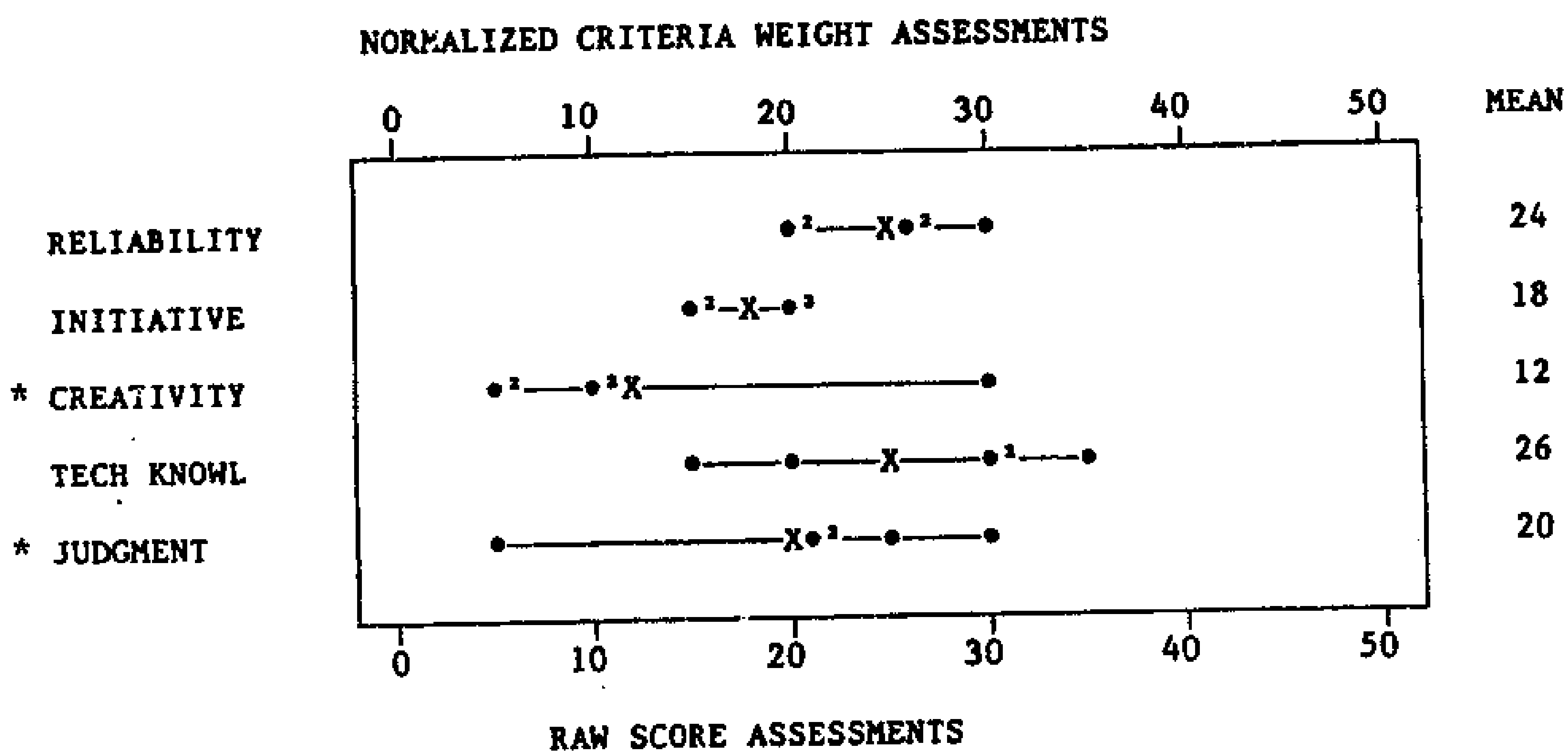
	RELIABILITY	INITIATIVE	CREATIVITY	TECHNICAL KNOWLEDGE	JUDGMENT	OVERALL
Earl	94	100	100	85	100	95
Dave	100	95	50	100	75	88
Frank	70	54	48	92	75	71
Baker	54	50	25	80	60	58
Able	48	30	10	70	0	36
George	0	0	20	30	20	14
Charlie	3	15	0	0	25	8

CRITERION WEIGHTS:

Job Performance	0.24	0.18	0.12	0.26	0.20

significant impact on overall results, thus permitting debate to be curtailed, since its outcome will have no impact on the final evaluation results.

FIGURE 42 -- GRAPHICAL SUMMARY OF NORMALIZED CRITERIA WEIGHTS



Superscript numbers indicate multiple weights assessed at that point. The asterisks at the left indicate greater than 25 point disagreement between the highest and the lowest weights assessed for that criterion. The X's indicate the algebraic mean of the individual weights received. The • show the individual weights.

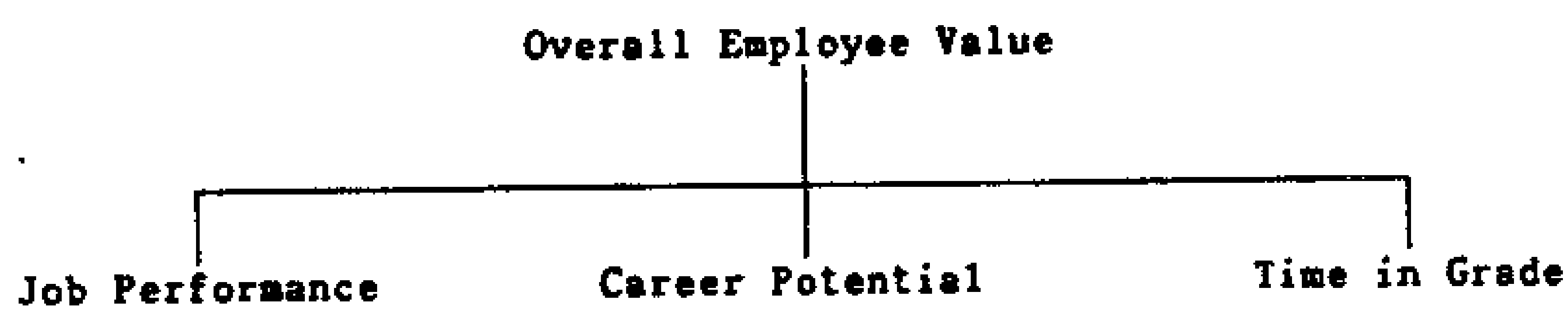
Extensions of this Evaluation Approach

The example presented is not fully complete, either in terms of the full range of activities involved in using a bottom-up, hierarchical rating scheme, or in terms of addressing all CSP evaluative needs.

In terms of the first point, no attention has been given to the verbal documentation procedures that should be incorporated into the evaluation process to make the audit trail more useful. The record of numerical assessments established by the procedure discussed in this paper is helpful, but verbal rational for the assessments need to be captured also in order to document not only the "what" of the assessments, but also the "why."

Also not addressed is the issue of how to compare people across job categories (e.g., how to compare a GS-15 supervisor to a GS-15 project officer). This can be done by extending the procedure outlined above: for example, by adding 1) an absolute assessment scale for each criterion, so that the relative assessments can (by assigning absolute scaled scores to the best and worst person in each category only) be converted into absolute scores, and 2) a hierarchical structure (perhaps like the one shown in Figure 43).

FIGURE 43 -- HIERARCHICAL STRUCTURE FOR COMPARISONS ACROSS JOB TYPES



An advantage, in addition to comparisons across GS levels, to extending the procedure to an absolute scale would be that CSP evaluations from one year could be compared to earlier years. This would permit tracking the progress of individuals and detecting assessment phenomena such as "grade creep." Also, the extension could provide an additional consistency check on the criteria weightings assessed, since the absolute-scaled criterion categories could be weighted as an independent exercise, even though the relative and the absolute weightings will have a fixed mathematical relationship imposed by the relative to absolute-scale score conversion.

RESOURCE ALLOCATION DESIGNS

Hierarchical and utility based scaling procedures can be used as resource allocation aids. An example evaluation structure is described in tutorial fashion below.

Resource allocation involves distributing limited resources in a manner that most effectively accomplishes assigned goals and missions. Usually, sufficient resources are not available to accomplish all assigned (or desired) tasks. A common example is deciding how to task an analytical staff to perform the intelligence studies that are needed. Decisions like this staffing problem have many similarities to building a factory or selecting the set of appliances to buy for a new home. For the appliance selection case, one must decide how much to invest into each of several categories of equipment (e.g., how much to spend on a refrigerator, a freezer, a dishwasher, an air-conditioner, etc.). Each category of appliance could be funded at any number of levels, depending on the number of models you've limited your selection to. Some categories might be optional (a trash compactor), whereas others (e.g., a refrigerator) require investment of at least the cost of the cheapest available model. The goal in choosing the appliance package that is to be purchased is to obtain the maximum utility from the appliances for the investment made. If we have an unlimited budget, then it is an easy matter to maximize the worth of the package selected: simply buy the best of everything. This will, also maximize our expense (assuming for the moment that the best in each category is also the most costly). If our budget is limited, our attempt to maximize utility becomes more difficult as we try to assess the relative benefits of investing the whatever extra capital we have (that money in the budget that exceeds the cost of buying the cheapest of every necessary appliance) in one appliance type versus one (or more) other type.

Investing analytical man-years in various categories of intelligence project areas is highly analogous. Our budget is limited to the number of analysts in our employ. The utility we hope to maximize is the policy support capability that results from the slotting allocations chosen. Our "purchase" options consist of the number of ways we consider it feasible to staff a given analytical effort (rather than the set of appliance models offered by some external player).

The basic task is to

a) Determine the set of feasible staffing options for each intelligence topic being considered for coverage,

b) Determine the relative benefit of each of the potential options within a given intelligence project area,

c) Determine the relative importance of the various intelligence topics under consideration.

The structure developed below should help to ensure that each of these issues is addressed explicitly during the decision process. It therefore should provide the benefit of acting as a communication aid to the members of the decision group. It also provides an audit trail to the decision process, and is easy to document with the rational used in making the assessments.

For a illustrative example, let's assume that we want to assign analysts to cover seven countries: the USSR, PRC, France, Cuba, Canada, Argentina, and Iraq. We first need to define the "buy" options available for each country. Let's assume, for the purpose of keeping this illustration simple, that we feel that there are seven reasonable ways that the USSR could be staffed: with 3, 4, 5, 7, 10, 12, or 15 analysts.[43] Assume

----

[43] These numbers imply that the USSR will be staffed <u>at a minimum</u> with 3 analysts. That is the worst that can happen; so (as will be discussed

also, that we feel that Canada could be staffed at any of three staffing strengths: 1/4, 1/2, or 1 analyst (where 1/4 and 1/2 mean that some analyst is spending 25% or 50% of his or her time on Canada and the remaining time on another country(s). Graphically, we can represent these proposed options (along with options for the other seven countries as is shown in Figure ** The boxes represent the options and the numbers in the lower

FIGURE 44 -- PROPOSED STAFFING OPTIONS

USSR	3	4	5	7	10	12	15
PRC	2	3	5	7			
France	.5	1	2				
Cuba	.2	.5	1	2	3		
Canada	.25	.5	1				
Argentina	.5	1	2				
Iraq	1	2	3	5			

right half of each box, the "cost" in analysts to chose that option.

----

below) nothing is gai<u>ned</u> by choosing this option.

PROBLEM SOLVING TECHNIQUES
Assessing and Predicting Decisions

The information contained in Figure 44 represents a concise statement of an analysis of staffing options and their relative costs. The next step is to indicate the relative value of each of these options to the decision maker.[44] This is done in stages. The first stage involves comparing the relative worth of the options within a given country to the other options of the same country only. Since it will be done on a relative basis, the limits of the numerical evaluation scale to be used can be set arbitrarily. It is convenient to use 0 (zero) to 100. Although numbers will be assigned subjectively, it will be done in a manner that will permit the numbers to be manipulated as interval-scaled data. Consider the USSR as an example of how the options within each country would be assigned utility values.

First, note that there are seven option levels to evaluate. Let's assume that we are comfortable assigning a single value to each option, later we'll discuss how to do this if multiple, conflicting evaluation criteria make this difficult; and also assume, as usually is the case, that the most expensive option is worth the most and the least expensive is worth the least to us. Based on these assumptions, we can assign the 15-analyst option a 100 (full-worth) and the 3-analyst option a zero. (The zero here does not imply that the 3-analyst option has no worth at all, but rather that it has no utility to us since it is the worst that we can do from the available USSR choices. That is, nothing is gained by choosing this option. The method to be used later to evaluate the relative importance of the differing countries will take this specific use of the zero utility assignment into account.)

Having assigned the 100 and the 0 to the best and worst USSR options, we now must assess the relative "goodness" of the intermediate options. To do this, we compare them to the best and worst options by considering the difference in value between the best and worst options. Assume that we feel that the 7-man option (which adds 4 people to the lowest, 3-man option) can provide about half of the additional capability that we would obtain by adding 12 more people to obtain the 15-man (highest) option. In this case we would assign a value of 50 (half-way between the zero and 100 extremes of the best and worst options) to the 7-man staffing option. Values for all options for each country need to be assigned in an analogous fashion. (The value assignment process is conceptually simple, but is time consuming and does require frequent consistency checks to ensure that the values assigned can be treated as interval-scale data.)

Let's assume that we've assigned the utilities to each level that are indicated in the upper-left portion of each square shown in Figure 45.

---

[44] The task of assigning values to the options is not trivial or easy. The process will be treated rather cavalierly here to reduce the complexity and thereby minimize distraction from the cost-benefit structuring procedure being presented. Frequently the value (utility or regret) assessments are made with the aid of a hierarchical evaluation structure such as discussed on page 75.

FIGURE 45 -- STAFFING OPTION VALUES

USSR	0 / 3	15 / 4	35 / 5	50 / 7	70 / 10	85 / 12	100/ 15
PRC	0 / 2	35 / 3	75 / 5	100/ 7			
France	0 / .5	45 / 1	100/ 2				
Cuba	0 / .2	20 / .5	50 / 1	95 / 2	100/ 3		
Canada	0 / .25	85 / .5	100/ 1				
Argentina	0 / .5	70 / 1	100/ 2				
Iraq	0 / 1	20 / 2	50 / 3	100/ 5			

The utility assignments that we just completed for each of the staffing levels provide a comparison only within a given country. We now must provide a means for comparing the meaning the USSR 100 to the Canada 100. This we do by weighting each of the countries in a manner similar (but not quite identical) to the method just used to assign values to the the different staffing levels within each country. We will weight the importance of each country by assessing the relative importance of the difference in capability between the worst and best staffing options of each country. We will assign an importance weight of 100 to the country which you assess to have the biggest difference between extreme staffing levels (most gain going from the smallest to the highest -- or most loss of opportunity going from the highest to the lowest levels). The major change in the way we assign values to the countries from the method used for staffing levels within countries is that (unless there is absolutely no improvement expected by going from the lowest to the highest staffing level within a country) no country needs to have a zero assigned to it. For the time being, let's assume the country weightings indicated in the right-hand

FIGURE 46 -- WEIGHTED STAFFING OPTIONS

USSR	0 / 3	15 / 4	35 / 5	50 / 7	70 / 10	85 / 12	100/ 15	100
PRC	0 / 2	35 / 3	75 / 5	100/ 7				80
France	0 / .5	45 / 1	100/ 2					20
Cuba	0 / .2	20 / .5	50 / 1	95 / 2	100/ 3			70
Canada	0 / .25	85 / .5	100/ 1					15
Argentina	0 / .5	70 / 1	100/ 2					25
Iraq	0 / 1	20 / 2	50 / 3	100/ 5				85

margin of Figure 46. The country weighting factors relate to the benefit values assigned to the staffing levels, not to the cost (people) values assigned to the levels. That is, the country weights tell us, for example that going from the lowest to the highest of the proposed staffing levels for Argentina will give us only 25% of the benefit that we get by going from the lowest to the highest staffing level for the USSR. Hence, the actual benefit of each Argentine value is only 25% of that shown in Figure 46.

ANALYSIS ACTIVITIES

Assessing and Predicting Decisions

Using the information developed above, we can calculate the cost-benefit ratio for each proposed staffing package chosen from our list. (The number of possible packages equals the number of possible combinations of one level chosen from each country. In our case, the number of possible combinations is

7x4x3x5x3x3x4 = 15,120

Each one of these packages has a specific cost-benefit ratio that can be computed by using the scaling factors assigned above and the marginal cost incurred by choosing any of the enhanced levels.[47] Fortunately, we do not need to be concerned with all of the 15,000-plus options. This can be seen by graphing the set of options using cost (in people) on the X axis and percent benefit (in terms of the total possible) on the Y axis. If we chose the cheapest option from each country, we obtain no (i.e., 0%) benefit, but it still costs us 7.45 people. If we chose the highest staffing level from each country, we get 100% of the possible benefit and "spend" 35 people. The graph covering all possible options will thus have a Y axis ranging from 0 to 100% benefit (marginal benefit) and an X axis that ranges from 7.45 to 35 people. The two packages just described will occupy the lower left and upper right (extreme) corners of the graph. The other 15,000-plus options, it turns out, will occupy points that fall into a roughly football shaped region that extends between these two extreme points. Figure 47 illustrates a typical example of what the scatter diagram of all possible "buy packages" might look.

The only possible packages that we need be concerned about (based on the numerical assessments) are the ones represented by points lying along the top edge of the football-shaped region of Figure 47. For example, assume that someone proposes the package that corresponds to the point in Figure 47 that is labeled by the P. If we like this package, then we should be even happier with the package labeled B since it gives us more benefit

---

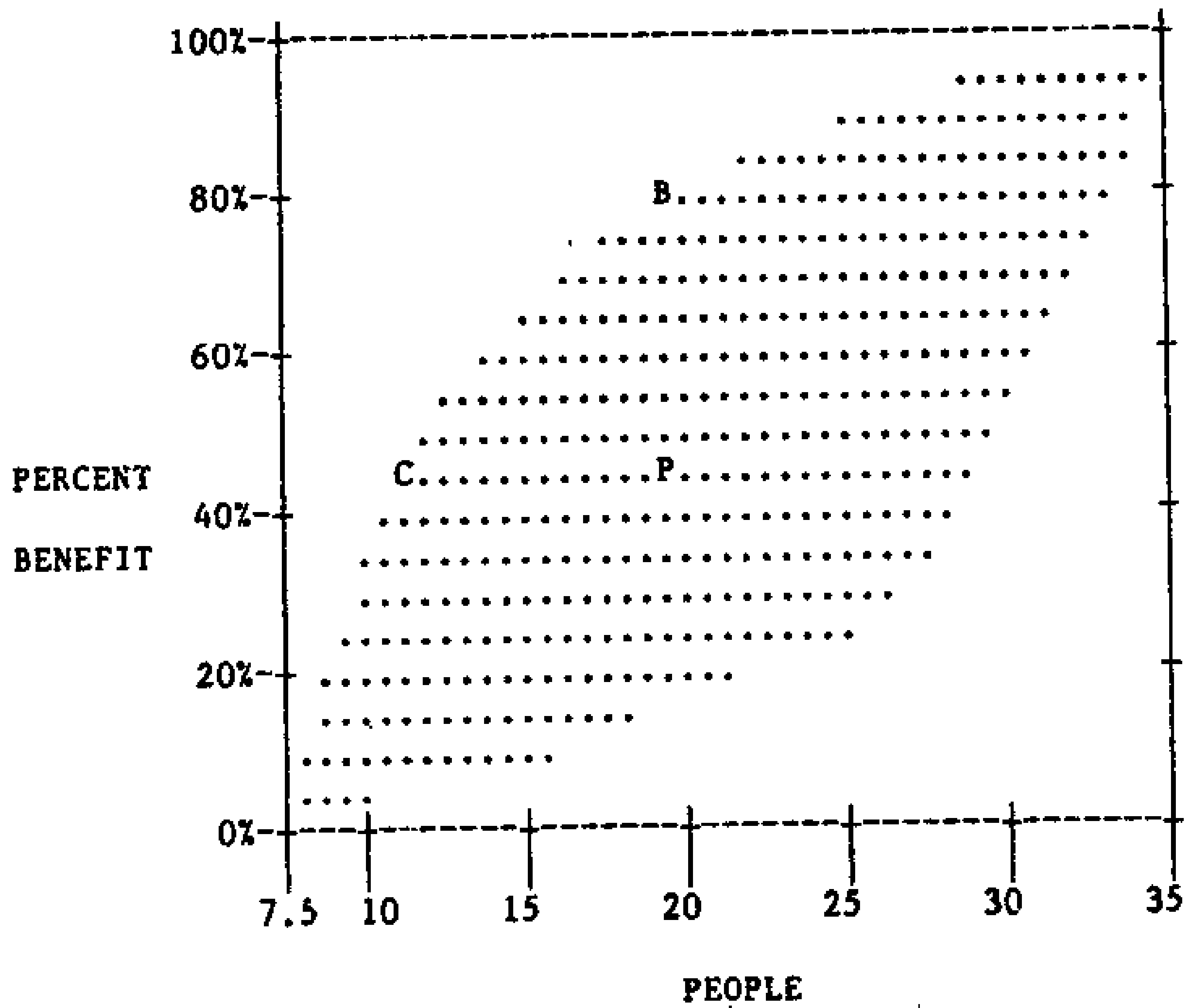
[47] Although it in fact costs people to man the lowest selection from each of the country staffing options, our definition of the lowest level for each country tells us that we will assign at least the specified number of analysts to that country. Therefore, cost-benefit calculations are based on the number of additional analysts that must be used to staff a more expensive option.

Thus, selecting the third USSR option along with the lowest option from all of the other countries would give us a buy package having a total benefit-cost ratio of 35/2 = 17.5 (the 35 benefit points assigned divided by the two additional people it costs to move from the lowest to the third option level). If we add the most expensive PRC option to this package, the benefit-cost ratio becomes:

$$(35 + (100 * 0.25)) / (2 + 5) = (35 + 25) / 7 = 60/7 = 8.57$$

since we still have the 35 benefit points from the USSR option and we gain another 25 points (100 times the one quarter scaling factor relative to the USSR) from the PRC, but must divide these 60 benefit points by the 7 people (2 USSR and 5 PRC) that we are spending to buy this more expensive package. Clearly, the first package has a better benefit-cost ratio (more benefit per person spent).

FIGURE 47 -- GRAPHICAL DISPLAY OF STAFFING OPTION RANGE



for the same cost, or with package C since it gives us the same benefit for far less cost.

For the specific application proposed above, namely, most efficiently allocating analysts, we would want to select the package corresponding to the highest point lying vertically above the number on the X axis that corresponds to the number of analysts we have available. Fortunately, computer software exists that will do all the necessary mathematics for us. The software also will permit us to do several other useful things:

a. We can have it define the optimum package for any chosen cost (e.g., 25 analysts) or benefit.

b. We can have it display all of the packages lying on the optimal "frontier" of the football shaped region. (Known in the literature as the Parado Optimal Frontier)

c. We can have it plot, in a Figure 47 type graph, where any specifically proposed package lies.

Needless to say, software options like these are useful for determining what would happen to our analytical capabilities if we were to be cut back 3 analysts to an austere level, or what extra we could do if we were given 5 additional analysts.

In addition to the country weighting method used above, weights can be assigned in ways that will take into account multiple, possibly conflicting criteria. (This method uses a hierarchical structure to assess utility values, in a manner similar to that discussed in the information evaluation section, page 75.) One available software package permits multiple criteria weights to be used. Another permits multi-level hierarchical structures to be evaluated.

GAME THEORY

INTRODUCTION

The theory of games analyzes mathematically a class of problems which might be called problems of strategy. The crux of game theory is the principle of choosing the strategy which minimizes the maximum expected loss (called minimax) or which maximizes the maximum possible gain (called maximax). Game theory defines a solution to a game as a set of imputations (payments) which satisfies this principle for all players.

Consider a game of tic-tac-toe. You know at any moment in the game what the moves available to your opponent are, but you do not know which one he will choose. The only information you have is that his choice will not, in general, be completely random; he will make a move which is designed in some way to increase his chance of winning and diminish yours. Thus, the situation is one of uncertainty rather than risk.

Your goals are similar to your opponent's. Your problem is: what strategy should you adopt? The theory of games offers rules about how to choose among the strategies available to yoo. In the case of tic-tac-toe, the rules are trivial; but in more complicated games of strategy, these rules may be useful.

PROBLEM SOLVING TECHNIQUES
Assessing and Predicting Decisions

TERMINOLOGY

To talk about game theory, a few technical terms are necessary. A strategy is a set of personal rules for playing the game. For each possible move on your part, your opponent will have a possible set of responses. For each possible response by your opponent, you will have a set of responses, and so on through the game. A strategy is a list which specifies what your move will be for every conceivable previous set of moves of the particular game you are playing.

Associated with strategies are imputations. An imputation is a set of payments made as a result of a game, one to each player. In general, different imputations will be associated with different sets of strategies, but for any given set of strategies there may be more than one imputation (in games involving coalitions).

Imputation X is said to dominate imputation Y if one or more of the players has separately greater gains (or smaller losses) in X than in Y and can, by acting together (in the case of more than one player), enforce the occurrence of X, or of some other imputation at least as good. The relationship of domination is not transitive.

A solution is a set of imputations, none of which dominates another, such that every imputation outside the solution is dominated by at least one imputation within the solution. The task of Game Theory is to find solutions. For any game, there may be one or more than one.

The payoff matrix presents your strategies, your opponent's strategies, and the values of the outcomes (represented at the matrix intersections) from both your perspective and that of your opponent.

THE MINIMAX LOSS PRINCIPLE

The notions of domination and of solution imply a new fundamental rule for decision making. This rule is the rule of minimizing the maximum loss, or, more briefly, minimax loss. In other words, the rule is to consider, for each possible strategy that you could adopt, what the worst possible outcome is, and then to select that strategy which would have the least ill-effects if the worst possible outcome happened. Another way of putting the same idea is to call it the principle of maximizing the minimum gain, or maximum gain. This rule makes considerable sense in two-person games when you consider that the other player is out to get you, and so will do his best to make the worst possible outcome for you occur. If this rule is expressed geometrically, it asserts that the point you should seek is a saddle-point, like the highest point in a mountain pass (the best rule for crossing mountains is to minimize the maximum height, so explorers seek out such saddle-points).

ANALYSIS ACTIVITIES
Assessing and Predicting Decisions

Before we go any further, we need a few more definitions. Games may be among any number of players, but the simplest game is a two-person game, and it is this kind of game which has been most extensively and most successfully analyzed. Fundamentally, two kinds of payoff arrangements are possible. The simplest and most common is the one in which one player wins what the other player loses, or, more generally, the one for which the sum of all the payments made as a result of the game is zero. This is called a zero-sum game. In nonzero-sum games, analytical complexities arise. These can be diminished by assuming the existence of a fictitious extra player, who wins or loses enough to bring the sum of payments back to zero. Such a fictitious player cannot be assumed to have a strategy and cannot, of course, interact with any of the other players.

In zero-sum, two-person games, each player (according to Game Theory) should pick his minimax strategy. But will this result in a stable solution? Not always. Sometimes the surface representing the possible outcomes of the game does not have a saddle-point. In this case, if player A chooses his minimax strategy, then player B will have an incentive not to use his own minimax strategy, because having found out his opponent's strategy, he can gain more by some other strategy. Thus, the game has no solution.

Various resolutions of this problem are possible. One is the notion of a mixed strategy, which is a probability distribution of two or more pure strategies. The fundamental theorem of the theory of games is that if both players in a zero-sum two-person game adopt mixed strategies which minimize expected loss, then the game will always have a saddle-point. Thus, each person will get, in the long run, his expected loss and will have no incentive to change his behavior even if he should discover what his opponent's mixed strategy is. Since A is already getting the minimum possible under the strategy he chose, any change in strategy by B will only increase A's payoff, and therefore cause B to gain less or lose more than he would by his own minimax strategy. The same is true of B.

Games involving more than two people introduce a new element--the possibility that two or more players will cooperate to beat the rest. Such a cooperative agreement is called a coalition, and it frequently involves side-payments among members of the coalition. The method of analysis for three-or-more-person games is to consider all possible coalitions and to solve the game for each coalition on the principles of a two-person game. This works fairly well for three-person games, but gets more complicated and less satisfactory for still more people.

TWO-PERSON ZERO-SUM GAMES

Two-person, zero-sum games that situations of conflict where there are only two participants. The success of one is at the expense of the other. Each player selects and executes those strategies and tactics which he believes will result in his winning the game. Experience often enables each player to predict with some accuracy the reactions of his opponent to some particular strategy he might pursue. (Many management situations, obviously, involve the participation of many persons and thus are not examples of two-person games. The mathematics for three-person and larger games, however, is complex. Just remember that the basics underlying the generation of optimum or winning strategies in conflict situations must respect the same principles of inductive and deductive logic, regardless of the number of participants.)

A two-person game is illustrated in Table 1. The players, X and Y, are equal in intelligence and ability. Each has a choice of two strategies. Each knows the outcomes (referred to as payoffs) for every possible combination of strategies; these are shown in the body of the table. Note that the game is biased against player Y; but since he is required to play, he will do his best. This is roughly equivalent to the business situation in which short-run loss is inevitable; these losses must be minimized by good strategy. The solution to this simple game is easily obtained by analyzing the possible strategies of each player:

Table 1

1. X wins the game only by playing his strategy M; thus he plays M all the time.

2. Y realizes that X will play strategy M all the time and, in an effort to minimize X's gains, plays his strategy Q.

3. The solution to the game is thus M, Q (strategy M and strategy Q).

4. X wins 2 points (Y loses 2 points) each time the game is played; thus the value of the game to X is 2; the value of the game to Y is -2.

The term value of the game used in this sense is the average winnings per play over a long series of plays. Though player Y loses in this game, he is still playing his optimum strategy, i.e., minimizing his losses. If he had adopted strategy R, his losses would have been 3 points per play, on the average.

Standard Language for Games

The <u>payoff</u> <u>matrix</u> can be used to describe games in a much more concise form than that used in Table 1. Here is the same game in abbreviated form:

		Y	
X		2	3
		-2	-3

ANALYSIS ACTIVITIES

Assessing and Predicting Decisions

The game has now been expressed as a matrix. This matrix is called the payoff or payout matrix. The four individual payoff possibilities are expressed as numbers; a positive number indicates a payoff to the player who plays the rows (X); a negative number indicates a payoff to the player who plays the columns (Y). The possible strategies for player X are row 1 and row 2 (reading from top to bottom as in standard matrix form); the possible strategies for Y are column 1 and column 2 (reading from left to right as in standard matrix form).

The matrices below show examples of several other two-person zero-sum games. Note that more than two strategies can be available to either or both of the players. In each case the value of the game is underlined. It represents the optimum strategy for each player; which for these type of games can be determined by ordinary algebra.

	Y	
X	2	4
	1	-3

		Y	
	3	2	-2
X	1	-3	-4
	0	1	-3

		Y	
X	2	0	4
	1	-3	2

	Y	
	1	-3
X	4	0
	3	-1

	Y	
	1	-3
X	2	4
	-1	5

Pure Strategies and Saddle Points

In each of the cases above, there is one strategy for player X and one strategy for player Y that will eventually be played each time. They may experiment for a while, but in time they will adopt the strategy we have illustrated; this assumes, of course, that each player desires to win (or to minimize his losses if he cannot win). In each of these games, each player has a pure strategy, one he plays all the time. The payoff which is obtained when each player plays his pure strategy is called a saddle point; or expressed a little differently, the saddle point is the value of a game in which each player has a pure strategy.

A saddle point can be recognized because it is both the smallest numerical value in its column. Ponder for a moment the significance of this. Player Y would rather have as a payoff the smallest numerical value in any column. Naturally, when there is one numerical value which satisfies both these conditions (a saddle point), both players will be playing optimally if each choose that value.

Mixed Strategies

Not all two-person games have a saddle point. Consider the following games:

1	0
-4	3

4	-2	3
0	5	6

1	2
3	4
5	-6

1	6	8
2	7	-9
-4	9	12

In the cases where no saddle point exists, the players resort to a mixed strategy: To optimize his winnings, player X will play each of his rows a certain part of the time, and player Y will play each of his columns a certain part of the time. X must determine what proportion off the time he should play each row, and Y must determine what proportion of the time the should play each column.

Following is a simple two-person zero-sum game with no saddle point.

	Y	
X	5	1
	3	4

Our tasks are to determine what portion of his time player X should spend playing each of his rows and what portion of the time player Y should spend playing each of his columns. Since we are dealing in portions of time, we need to differentiate between the time player X spends playing row 1 and the time he spends playing row 2. The same type of breakdown must be made for player Y.

Let Q equal the fractional proportion of the time player X spends playing the first row; then 1 - Q equals the proportion of the time he would spend playing the second row. (For example if player X plays the first row 3/4 of the time, then 1-3/4 or 1/4 must equal the time he spends playing the second row.) The same concept, of course, applies to player Y and his distribution of time between the rows and between the columns. This is illustrated below:

1. Player X plays the first row Q of the time (Q is between 0 and 100 per cent).

2. Player X plays the second row 10-Q of the time.

ANALYSIS ACTIVITIES
Assessing and Predicting Decisions

3. Player Y plays the first column P% of the time (P is between 0 and 100 per cent).

4. Player Y plays the second column 100-P% of the time.

To solve for the unknown fractions P and Q, consider player X first. Logically, X wants to divide his plays between his rows so that his expected winnings from playing the first row will exactly equal his expected winnings from playing the second row--no matter what Y does. (Of course, X's opponent Y is assumed to be intelligent and is expected to adopt his optimum strategy, X would reason out his strategy differently; he would simply look for the obvious loopholes in Y's strategy and play accordingly.)

	If Y plays column 1	If Y plays column 2
X plays row 1 Q of the time	X wins 5 points Q of the time	X wins 1 point Q of the time
X plays row 2 (1 - Q) of the time	X wins 3 points (1 - Q) of the time	X wins 4 points (1 - Q) of the time
X's expected winnings	5Q + 3(1 - Q) when Y plays column 1	1Q + 4(1 - Q) when Y plays column 2

To make X's expected winnings when Y plays column 1 equal to X's expected winnings when Y plays column 2, we let $5Q + 3(1-Q)$ equal $1Q + 4(1-Q)$ and solve for the one particular Q which does equate the two expectations:

$$5Q + 3(1 - Q) = 1Q + 4(1 - Q)$$
$$5Q + 3 - 3Q = 1Q + 4 - 4Q$$
$$5Q = 1$$
$$Q = 1/5$$

Therefore, $1 - Q = 4/5$.

Our algebraic solution thus indicates that player X plays the first row 1/5 of the time and the second row 4/5 of the time.

To solve for player Y's optimal strategies use the same algebraic method. Player Y wants to divide his time between his columns so that no matter what X does about the rows, Y will maximize his winnings (minimize his losses) over time. Y's choice of strategies between the columns can be represented in algebraic form as shown below.

Equate Y's expected losses when X plays row 1 with Y's expected losses when X plays row 2:

$$5P + 1(1 - P) = 3P + 4(1 - P)$$

and solve for P:

$$5P + 1 - P = 3P + 4 - 4P$$
$$4P + 1 = 4 - P$$
$$5P = 3$$
$$P = 3/5$$

Therefore, $1 - P = 2/5$.

Our solution for Y's best strategy is for him to play the first column 3/5 of the time and the second column 2/5 of the time. Of course, when we indicate that a player should play each of his choices a certain proportion of the time, realize that this division of time between the rows or columns must be done at random, without any discernible pattern. If one of the players begins to notice a pattern in the play of his opponent--if X notices that Y plays column 1 three times and then column 2 twice, over and over again in that order--he should adjust his strategy to take advantage of Y's disclosure of his future pattern. On the other hand, if both players play their best strategies but without a discernible pattern, then the strategies we have determined represent the best possible divisions of time between the rows or columns.

Now that we have solved for the optimum mixed strategies, we can calculate the value of this game. Our original game, together with the optimum strategies for each of the players, is presented below:

		Y	
		3/5	2/5
X	1/5	5	1
	4/5	3	4

Look at the game from player X's point of view, we can reason as follows:

1. During the 3/5 of the time that Y plays column 1, X wins 5 points 1/5 of the time and 3 points 4/5 of the time.

2. During the 2/5 of the time that Y plays column 2, X wins 1 point 1/5 of the time and 4 points 4/5 of the time.

Therefore, X's total expected winnings over time are the sum of statements 1 and 2 above:

1		2
3/5 [5(1/5) + 3(4/5)]	+	2/5 [1(1/5) + 4(4/5)]
3/5 (5/5 + 12/5)	+	2/5 (1/5 + 16/5)
3/5 (17/5)	+	2/5 (17/5)

= 17/5 = Value of the Game

This means that player X, if he plays his optimal strategies, can expect to win an average of 3 2/5 points for each play of the game. From our earlier observation, we know that X will be the winner of this game, since the value is a positive number. If the value of the game had been a negative number, Y would have been the winner. This, of course, could not be true in this particular game, since the game was slanted in X's favor in that it contained no negative payoffs in the original matrix.

We could have arrived at the same value of the game looking at it from Y's point of view:

1. During the 1/5 of the time that X plays row 1, Y loses 5 points 3/5 of the time and 1 point 2/5 of the time.

2. During the 4/5 of the time that X plays row 2, Y loses 3 points 3/5 of the time and 4 points 2/5 of the time.

Therefore, Y's total expected losses over time are the sum of statements 1 and 2:

1		2
1/5 [5(3/5) + 1(2/5)	+	4/5[3(3/5) + 4(2/5)]
1/5 (15/5 + 2/5)	+	4/5 (9/5 + 8/5)
1/5 (27/5)	+	4/5 (17/5)

= 17/5

Again we see that the value of the game is 3 2/5; since it is a positive number, we know that X wins. The term "value of the game" does not mean that X will win 3 2/5 points each time these two players play; it means instead that X's winnings over many plays of the game will average 3 2/5 points per game.

STATISTICAL DECISIONS

By viewing the problem of statistical decision making as one of playing a game against Nature, the statistician must decide, on the basis of observations (which cost something to make) between policies, each of which has a possible gain or loss. In some cases, all these gains and losses and the cost of observing can be exactly calculated, as in industrial quality control. In other cases, as in theoretical research, it is necessary to make some assumption about the cost of being wrong and the gain of being right. At any rate, when they are put in this form, it is obvious that the ingredients of the problem of statistical decision making have a game-like sound.

A very frequent criticism of the minimax approach to games against Nature is that Nature is not hostile, as is the opponent in a two-person game. Nature will not, in general, use a minimax strategy. But, even though Nature is not hostile, she does not offer any way of assigning a probability to each possible outcome. In other words, statistical decision making is a problem of uncertainty, rather than of risk.

Other suggested principles are: maximizing the maximum expected gain, and maximizing some weighted average of the maximum and minimum expected gains. None of these principles commands general acceptance; each can be made to show peculiar consequences under some conditions.

ANALYSIS ACTIVITIES
Biases in Evaluating Information

INTERPRETING ANALYTICAL RESULTS

BIASES IN EVALUATION OF EVIDENCE

Article by Dick Heuer, retired OPA analyst**

PREFACE

Evaluation of evidence is a crucial step in analysis, but what evidence we rely on and how we interpret it are influenced by a variety of extraneous factors. Information presented in vivid and concrete detail often has unwarranted impact, while we tend to disregard abstract or statistical information that may have greater evidential value. Data on prior probabilities are commonly ignored unless they illuminate causal relationships. We are probably not as good as we should be in taking the absence of evidence into account. We are also oversensitive to the consistency of the evidence, and insufficiently sensitive to the reliability of the evidence we handle. Finally, there is reason to believe that impressions often persevere even after the evidence on which they are based has been totally discredited.

THE ANALYTICAL ENVIRONMENT

The intelligence analyst works in a somewhat unique informational environment. Evidence comes from an unusually diverse set of sources: newspapers and wire sources, observations by American Embassy officers, reports from controlled agents and casual informants, information exchange with foreign governments, photoreconnaissance, and signals intelligence. Each source has its own unique strengths, weaknesses, potential or actual biases, and vulnerability to deception. The most salient characteristic of the information environment is its diversity -- multiple sources, each with varying degrees of reliability, and each commonly reporting information which by itself is incomplete and often inconsistent or even incompatible

---

** In this article, which appeared in ANALYTICAL METHODS REVIEW, A Review of Ideas and Applications, January 1981, Dick Heuer, who has written a series of articles on the psychology of intelligence analysis for the Office of Political Analysis, analyzes the myriad problems analysts face in evaluating evidence. (U) with reporting from other sources. Conflicting information of uncertain reliability is endemic to intelligence analysis, as is the need to make rapid judgments on current events even before all the evidence is in.

The analyst has only very limited control over the stream of information; tasking of sources to report on specific subjects is a cumbersome and commonly time-consuming process. Evidence on some important topics is sporadic or nonexistent. Almost all human source information is secondhand at best; because of limitations imposed by their own CIA employment, most intelligence analysts have spent less time in the country they are analyzing and have fewer contacts with nationals of that country than their academic and other government colleagues.

Because of the nature of this analytical environment, the intelligence analyst may be more vulnerable to some biases in the evaluation of evidence than analysts in many other domains. For example, the intelligence analyst is more likely to have to cope with evidence of questionable reliability, or to factor in the significance of the absence of any evidence at all. Similarly, the intelligence analyst is commonly so beset with conflicting evidence that consistency of evidence is likely to have any even stronger impact than in most other circumstances.

With this background, we turn now to discussion of specific biases. Most of the biases are unrelated to each other and are grouped together here because they all concern some aspect of the evaluation of evidence.

THE VIVIDNESS CRITERION

Information that is vivid, concrete, and personal has a greater impact on our thinking than pallid, abstract information that may actually have substantially greater value as evidence. This fact is well understood by writers of advertising, who exploit it relentlessly. It is also well known to intelligence analysts, but we often fail to recognize circumstances in which this phenomenon appears and the extent to which it affects our thinking.

In this section, we argue that the impact of information on our thinking is only imperfectly related to its true value as evidence,(1) and we point out two circumstances of particular relevance to intelligence analysts when this may be the case:

-- Information that we perceive directly, that we hear with our own ears or see with our own eyes, is likely to have greater impact than information we receive second- hand that may have greater evidential value.

-- Case histories and anecdotes will have greater impact than more informative but abstract aggregate or statistical data.

ANALYSIS ACTIVITIES

Biases in Evaluating Information

Events we experience personally are more memorable than those we only read about. Concrete words are easier to remember than abstract words,(2) and words of all types are easier to recall than numbers. In short, vivid, concrete, personally relevant information is more likely to be stored and remembered than abstract reasoning or statistical summaries, and, therefore, is more likely to have continuing impact on our thinking in the future as well as a greater immediate impact. Such information is more likely to attract and hold our attention, and whatever we focus attention on is more likely to be perceived as having causal influence. The availability rule (3) also plays a role; information that is vivid comes to mind more readily and is therefore more available to be used in our thinking.

Intelligence analysts generally work with secondhand information; the information they receive is mediated by the written word of others rather than being perceived directly with their own eyes and ears. The occasions when an analyst visits the country whose affairs he or she is analyzing or speaks directly with a national from that country are exceptional and noteworthy. Such vivid experiences are often a source of new insights, but they can also be a cause of self-deception.

That concrete, sensory data do and should enjoy a certain priority when weighing evidence is well established. When an abstract theory or secondhand report is contradicted by personal observation, the latter properly prevails under most circumstances. There are number of popular adages that caution us to mistrust secondhand data: "Don't believe everything you read," "You can prove anything with statistics," "Seeing is believing," "I'm from Missouri...."

It is curious that we have no comparable maxims to warn us against being misled by our own observations. Seeing should not always be believing. Consider, for example, the reporters covering Senator George McGovern in the 1972 presidential campaign. On the eve of the election, they agreed unanimously that McGovern could not possibly lose by more 10 percentage points, even though they knew he was trailing by twenty points in all the polls and that no major poll had been wrong by more than 3 percent for twenty-four years. They had observed enthusiastic crowds cheering McGovern all over the country, and they gave disproportionate weight to these vivid personal observations despite the obvious bias in the sample of people being observed.(4)

Personal observations by intelligence analysts and agents can be as deceptive as secondhand accounts. Any single person, especially a visitor to a foreign country, normally becomes familiar with only a small sample of people representing a narrow segment of the total society. Incomplete and distorted perceptions are a common result.

A familiar form of this error is the single, vivid case that outweighs a much larger body of statistical evidence or conclusions reached by abstract reasoning. When a potential car buyer overhears a stranger complaining about how his Volvo turned out to be a lemon, this may have as much impact on the potential buyer's thinking as statistics in Consumer Report on the average annual repair costs for foreign-made cars. If the personal testimony comes from the potential buyer's brother or close friend, it is likely to be given even more weight. Yet the logical status of this new information is to increase by one the sample on which the Consumer Report statistics were based; the personal experience of a single Volvo owner has virtually no evidential value at all.

Nisbett and Ross label this the "man-who" syndrome and provide the following illustrations:

-- "But I know a man who smoked three packs of cigarettes a day and lived to be ninety-nine."

-- "I've never been to Turkey but just last month I met a man who had, and he found it...."

Needless to say, a "man-who" example seldom merits the evidential weight intended by the person citing the example, or the weight often accorded to it by the recipient.

The most serious implication of vividness as a criterion that determines impact of evidence is that certain kinds of very valuable evidence will have little influence simply because they are abstract. Statistical data, in particular, lack the rich and concrete detail to evoke vivid images, and they are often overlooked, ignored, or minimized. For example:

The two decades since publication of the surgeon general's report linking cigarette smoking to cancer have witnessed no per-capita decline in cigarette consumption. One subpopulation, however, has shown a definite decline in smoking -- physicians. This exception, however, probably should not be attributed simply to the physician's superior knowledge and intelligence, or the desire to 'set a good example.' All physicians are aware of the statistical evidence, yet there are clear differences among their specialties in the degree of decline in smoking. The probability that a physician smokes is directly related to the distance of the physician's specialty from the lungs. Physicians who diagnose and treat lung cancer victims are quite unlikely to smoke, and radiologists have the very lowest rate of smoking. Informational vividness seems to influence even the most sophisticated people, even when they have been exposed to the most probative data.(5)

Personal anecdotes, actual accounts of people's responsiveness or indifference to information sources, as well as controlled laboratory experiments can be cited ad infinitum to illustrate the proposition that data summaries, despite their logically compelling implications, have less impact than does inferior but more vivid evidence.(6) It seems likely that intelligence analysts, too, assign insufficient weight to statistical information.

In summary, analysts should give little weight to anecdotes and personal case histories unless they are known to be typical, and perhaps no weight at all if aggregate data based on a more valid sample can be obtained.

ANALYSIS ACTIVITIES
Biases in Evaluating Information

BIAS FAVORING CAUSAL EVIDENCE

The bias in favor of causal evidence, with its various ramifications, is one of the more difficult biases to grasp. Because it is easier to illustrate than to define, we start with a simple experiment. The reader will have to exert some effort to follow the reasoning through this and several succeeding experiments.

A panel of personnel officers interviewed a group of applicants to fill a middle-management job opening. The group of interviewees consisted of 30 engineers and 70 lawyers. The personnel officers summarized their impressions of each. Please answer the following questions, actually writing down your numerical probability estimate on a separate paper.(7)

1. Jack is 45 years old. He is married and has four children. He is generally conservative, careful, and ambitious. He shows no interest in political and social issues and spends most of his free time on his many hobbies which include home carpentry, sailing, and mathematical puzzles. What is the probability that Jack is one of the 30 engineers in the sample?

1. Dick is 30 years old. He is married with no children. A man of high ability and high motivation, he promises to be quite successful in his field. He is well liked by his colleagues. What is the probability that Dick is one of the 70 lawyers in the sample?

3. Harry is one of the men in the sample. What is the probability that Harry is an engineer rather than lawyer?

Typical answers to these questions are about 95 percent certainty that Jack is an engineer, a 50 percent probability that Dick is a lawyer, and a 30 percent probability that Harry is an engineer. In such inference tasks, people normally select the answer that is most representative of the evidence. The available evidence about Jack is very similar to our stereotype of an engineer, hence the high probability estimate that he is an engineer rather than a lawyer. The evidence about Dick, in question 2, does not help us distinguish between the two alternatives, hence the 50 percent estimate.

If you followed this common reasoning on the first two questions, the third question provides a clue that should have caused you to reevaluate your previous answers. Question 3 contains no descriptive information about Harry. In the absence of such evidence, we know only that of the 100 interviewees, 30 were engineers and 70 lawyers. Clearly, there is a 30 percent probability that Harry is an engineer. But is the distribution of engineers and lawyers in the sample not also relevant when answering the previous questions? Surely it is, for if there are more than twice as many lawyers as engineers, Jack may well be a lawyer despite the fact that he seems to fit the stereotype of an engineer. Personality and interests are useful predictors of occupational choice, but the relationship is not so strong as to warrant ignoring altogether what we call the 70 percent prior probability that Jack is a lawyer. For the question about Dick, in which the given evidence is irrelevant because it does not point to either engineer or lawyer, the proper answer is 70 percent rather than 50 percent. In the absence of valid evidence, predictions should be based on the prior probability as in Question 3.

In assessing any situation, an analyst often has two kinds of evidence available -- specific evidence about an individual case, and background information that summarizes or generalizes information about many similar cases. The case-specific evidence describes unique characteristics of the individual case. The background information, which is also referred to as prior probabilities, norms or base rates, reflects the fact that each individual case is more than just a collection of unique properties; it is also one member of a generic class of many cases that share common characteristics.

Jack, Dick and Harry are all individuals, each with his own peculiarities that may be more or less similar to the stereotyped characteristics we associate with engineers and lawyers. But they are also members of a group of individuals which has properties of its own -- the group of interviewees of which 30 were engineers and 70 lawyers. In making judgements about them one should use both the unique and the common information.

The optimal approach to the engineer-lawyer problem is to start with the assumption of a 30 percent probability that any single interviewee is an engineer, 70 percent that the interviewee is a lawyer. This is the prior probability before receipt of any case-specific information about Jack, Dick or Harry. The more confident we are that the interviewer's assessments are accurate,and the stronger the relationship assumed to exist between personality and interests and career the more weight we should assign to the case-specific evidence. Conversely, the lower our estimate of the reliability of an interviewer's assessments and the weaker the relationship between personality and career, the less we should deviate from the prior probability regardless of how much the evidence may seem to indicate that the interviewee is an engineer or a lawyer, as in Question 1 above. The proper answer to Question 1 is roughly 50 percent rather than 95 percent.

The key factor that explains this common error in reasoning is a strong preference for using causal evidence. We ignore the prior probabilities unless they elucidate some causal relationship that helps to explain the individual case. In Question 1, the case-specific data on Jack's hobbies and lack of interest in political and social issues are interpreted as part of a personality pattern that might cause him to select a career in engineering rather than law; they are, therefore, considered highly relevant. The fact that 70 percent of the interviewees were lawyers cannot cause Jack to be a lawyer, so this is commonly -- although erroneously -- perceived as irrelevant to a judgment about Jack's profession. The tendency to ignore non-causal background information is so strong that such information is not used even though the case-specific evidence is worthless, as in Question 2. It is only when case-specific causal evidence is totally lacking, as in Question 3, that people recognize the relevance of prior probabilities and use them in an appropriate manner.

ANALYSIS ACTIVITIES
Biases in Evaluating Information

We turn now to another experiment that illustrates the difference between causal and noncausal background data. This experiment also moves us closer to the kinds of judgments that might be made by intelligence analysts.(8)

During the Vietnam War, a fighter plane made a non-fatal strafing attack on a US reconnaissance mission at twilight. Both Cambodian and Vietnamese jets operate in the area. You know the following facts:

(a) 85 percent of the jet fighters in that area are Vietnamese, 15 percent are Cambodian.

(b) The pilot identified the fighter as Cambodian. The recognition capabilities of this pilot under appropriate visibility and flight conditions have been tested. When presented with a sample of fighters (half with Vietnamese markings and half with Cambodian) the pilot made correct identifications 80 percent of the time and erred 20 percent of the time.

Question: What is the probability that the fighter was Cambodian rather than Vietnamese?

A common procedure in answering this question is to reason as follows: We know the pilot identified the aircraft as Cambodian. We also know the pilot's identifications are correct 80 percent of the time; therefore, there is an 80 percent probability the fighter was Cambodian. This reasoning appears plausible but is incorrect, as is ignores the prior probability of 85 percent that the fighter was Vietnamese. It is actually more likely that the plane was Vietnamese than Cambodian despite the pilot's probably correct identification.

Readers who are unfamiliar with probabilistic reasoning and do not immediately grasp this point should imagine 100 cases in which the pilot has a similar encounter. Based on paragraph (a), we know that 85 of these encounters will be with Vietnamese aircraft 15 with Cambodian. Based on paragraph (b), we know that 80 percent or 68 of the 85 Vietnamese aircraft will be correctly identified as Vietnamese, while 20 percent or 17 will be incorrectly identified as Cambodian. Similarly, 80 percent or 12 of the 15 percent Cambodian aircraft will be correctly identified as Cambodian, while 20 percent or 3 will be incorrectly identified as Vietnamese. This makes a total of 71 Vietnamese and 29 Cambodian sightings, of which only 12 of the 29 Cambodian sightings are correct; the other 17 are incorrect sightings of Vietnamese aircraft. Therefore, when the pilot claims the attack was made by a Cambodian fighter, the probability that the craft was actually Cambodian is only 12/29ths or 41 percent, despite the fact that the pilot's identifications are correct 80 percent of the time.

This may seem like a mathematical trick, but it is not. The difference is due to the very strong prior probability of the pilot observing a Vietnamese aircraft. The difficulty in understanding this arises because our untrained intuitive judgment simply does not incorporate some of the basic statistical principles of probabilistic reasoning. We do not incorporate the prior probability into our reasoning because it does not seem relevant. And it does not seem relevant because there is no causal relationship between the background information on the percentages of jet fighters in the area and the pilot's observation.(9) The fact that 85 percent of the fighters in the area were Vietnamese and 15 percent Cambodian did not cause the attack to be made by a Cambodian rather than a Vietnamese.

To appreciate the different impact made by causally relevant background information, consider this alternative formulation of the same problem. In paragraph (a) of the problem, substitute the following:

> (a) Although the fighter forces of the two countries are roughly equal in number in this area, 85 percent of all harassment incidents involve Vietnamese fighters while 15 percent involve Cambodian fighters.

Although the problem remains mathematically and structurally the same, experiments with many test subjects show that it is quite different psychologically for it readily elicits a causal explanation relating the prior probabilities to the pilot's observation. If the Vietnamese have a propensity to harass and the Cambodians do not, we cannot ignore the prior probability that Vietnamese harassment is more probable than Cambodian. This linkage of the prior probability to a cause and effect relationship immediately raises the possibility that the pilot's observation was in error.

With this revised formulation of the problem, most people are likely to reason as follows: We know from past experience in cases like this that the harassment is usually conducted by Vietnamese aircraft. On the other hand, we have a fairly reliable report from our pilot that it was a Cambodian fighter. These two conflicting pieces of evidence cancel each other out; therefore, we do not know -- it is a roughly 50-50 situation whether it was Cambodian or Vietnamese. In employing this reasoning, we use the prior probability information, integrate it with the case-specific information, and arrive at a conclusion that is about as close to the optimal answer (still 41 percent) as one is going to get without using formal mathematical techniques.

It may be helpful to summarize these conclusions before moving on to discuss the problem in a more explicit intelligence context. The psychological impact of evidence is critically dependent upon its role in explaining a cause-effect relationship. Evidence in the form of background information, prior probabilities, base rates or norms that is not causally related to an outcome is generally ignored unless it is the only evidence on which to base a judgment. This is contrary to accepted principles of probabilistic reasoning. Background information that tells what the probabilities are prior to receiving specific evidence on the case at hand may be highly informative and should be used even if it does not illuminate cause and effect relationships. The more incomplete or unreliable the case-specific evidence is, or the more unpredictable the situation for whatever reason, the closer one's estimate should be to the prior

ANALYSIS ACTIVITIES
Biases in Evaluating Information

The point of this discussion is not simply that one has given an incorrect answer to one or more of these particular problems, but that by failing to use evidence concerning prior probabilities analysts may generally give less than optimal answers to all analytical problems of this generic type. This requires that we examine how common this type of problem is in life in general and in intelligence analysis in particular.

There are, of course, few problems in which prior, prob- abilities are explicitly given, as in the engineer/lawyer and Vietnamese/Cambodian aircraft experiments. They were given in these experiments to illustrate that even under the best of circumstances, when prior probabilities are presented as part of the problem, they are nevertheless still ignored unless they have some causal relationship to the outcome. When prior probabilities are not obvious but must be inferred or researched, they are even less likely to be used.

The so-called planning fallacy is a good example of a problem in which prior probabilities are not given in numerical terms but must be abstracted from experience. In planning a research project, I may estimate being able to complete it in four weeks. This estimate is based on relevant case-specific evidence: desired length of report, availability of source materials, difficulty of the subject matter, allowance for both predictable and unforeseeable interruptions, and so on. I also possess a body of experience with similar estimates I have made in the past: I, like many others, almost never complete a research project within the initially estimated time frame; I am lured by the immediacy and persuasiveness of the case-specific evidence; all the causally relevant evidence about the project indicates I should be able to complete the work in the time I have alloted for it. Even though I know from experience that this never happens, I do not learn from this experience. I continue to ignore the non-causal, probabilistic evidence based on many similar projects in the past, and to estimate completion dates that I hardly ever meet.

In the three examples cited so far, there is a clearly defined population of relevant cases --the 100 interviewees, all the hostile aircraft in the area, all the cases in which I have estimated the time required to complete a research project. In each of these examples, the people, objects or events have properties that are sufficiently similar that one can deduce properties and probabilities concerning individual cases from knowledge of the group as a whole.

For some questions of analytical interest, however, there is no set of sufficiently comparable cases. Consider an estimate of the outcome of events in Yugoslavia after the death of Tito. What is the relevant population of cases to be considered in order to estimate the prior probability of Soviet intervention, or of fragmentation of Yugoslavia into several smaller states? One could look at all instances of succession in communist systems but this would move one to a level of generalization that has little practical value for estimating the outcome of events in Yugoslavia. Clearly, there are a great many circumstances in intelligence analysis when it is not feasible to define a group of similar cases and estimate prior probabilities or norms.

There are, however, also many questions for which relevant prior probabilities are available. Sometimes the set of relevant prior cases is quite clear. For example, examination of earlier succession struggles in the Soviet Union provides insights (prior probabilities) concerning what we might expect during the succession to Brezhnev. In many other instances, the situation is more ambiguous. Should the prior probability of Israel bowing to US displeasure over its West Bank settlements policy be determined by studying all previous instances of US-Israeli disagreement, only previous Israeli responses to US initiatives concerning the West Bank, or all instances in which the US tries to influence the policies of friendly countries? This illustrates that there often are no generally accepted criteria for defining the relevant group of cases to be examined to establish a prior probability. The outcome of analysis often depends upon the set of comparable cases one chooses to analyze.

Readers trained in scientific method will recognize that what we are now referring to as prior probability or norms or relevant background data is what we defined earlier, in a different context,(10) as theory. In some cases, the identification of a prior probability is identical with the development of a theory that describes the probable relationship between variables in a set of comparable cases. Similarly, the use of case-specific data may be identical with the application of situational logic.

It is important to distinguish between prior probabilities based on prior cases occurring within the same country and those based on a set of cases occurring in diverse countries. Norms developed by studying a set of similar cases within the same country are seldom ignored by analysts; on the contrary, analysts rely on such evidence because there is a perceived causal connection. The underlying process in each case is assumed to be very similar, so that the same processes and forces that caused previous outcomes are anticipated to also effect the outcome of the present situation. In fact, the analyst examines these relevant historical precedents in order to identify the causal factors that may be most relevant in this generic type of situation in that country.

Comparison across countries serves the same logical function as comparison of historical precedents within the same country, but the psychological impact of such comparison is different because the causal connection is more remote or not apparent at all. Many phenomena have norms or prior probabilities that are relevant regardless of the country in which the phenomena occur. Collecting data and making generalizations that apply to generic phenomena occurring in many countries is what distinguishes the social scientific approach to research from the area studies approach. Area study is generally country-oriented, while social science compares phenomena across countries. This difference is shown in Figure 48 in schematic form. The vertical mode of analysis is our standard operating procedure when we examine case-specific evidence or evidence from a set of similar cases within the same country. It is what we do rather well, and the approach that is fostered by the fact that our own and our consumer's offices are necessarily organized by country and geographic region. For a variety of reasons, the horizontal, cross-country mode of analysis is much more difficult for us to do, and we do not do it often enough or well enough. Moreover, we often fail to use the results of such research even when it is available in academic studies. The reason is a failure to recognize that prior probabilities are relevant even though they do not reveal immediate causal relationships.

FIGURE 48 -- AREA STUDIES ORGANIZATION

Country	Country	Country	Country
Problem	Problem	Problem	Problem
Problem	Problem	Problem	Problem
Problem	Problem	Problem	Problem
Problem	Problem	Problem	Problem
Problem	Problem	Problem	Problem

If one examines the future prospects for the Saudi monarchy from a strictly Saudi viewpoint, there seems little prospect of the monarchical system being overthrown. If, however, one focuses on the generic phenomenon of social and political change in feudal societies suddenly exposed to a massive influx of foreigners and foreign ideas, the only relevant question is when, not if, the Saudi monarchy will undergo traumatic times. The case-specific evidence and prior probabilities point to entirely different conclusions.

The tendency to discount or ignore information concerning prior probabilities is enhanced by anything that increases the perception that the problem is unique, that it is a one-of-a-kind situation for which precedents in other countries are not applicable. Such perceptions of uniqueness are prompted by detailed acquaintance with a problem and by intense involvement with it. This is why the analysts who are most knowledgeable of and involved with affairs in any given country tend to use primarily case-specific data, while other observers are often guided by prior probabilities.

The country analyst is likely to believe that the policymaker or other generalist who lacks detailed familiarity with the case-specific evidence is poorly informed and therefore wrong. The policymaker, on the other hand, suspects that the country analyst has lost sight of the forest because of all the trees. Actually, it is not a question of which approach is right and which is wrong. Both are right as far as they go. The principal point is that both perspectives are necessary, that the best possible assessment of the present and forecast of the future can be made only by integrating both perspectives. Understanding the nature of societal change, for example, gives the Saudi analyst indicators to look for to seek the beginnings of a familiar process with a generally predictable outcome, and it permits the analyst to look beyond the limitations of the currently available evidence to forecast probable events for which the visible manifestations may not yet have begun to appear.

There are many similar situations in which cross-national research identifies relevant prior probabilities or norms. Often, it is not so much the norms themselves that are analytically important, as identification of the processes that underlie the outcomes in a number of similar cases. If Quebec separatism is examined only in a Canadian context, for example, driving forces behind this phenomenon may be overlooked. The resurgence of Scottish, Welsh and Corsican nationalism and of the long-dormant Occitan culture in southern France, exacerbation of ethnic conflict in Belgium, and even such developments as the American Indian Movement in the US and the secession of the French-speaking segment of the predominantly German-speaking Bern canton in Switzerland, indicate that assertion of national or ethnic identity has increased sharply in recent years throughout Western Europe and North America. The Quebec situation will unfold in a uniquely Canadian context, but we can understand the underlying forces at work only by a complementary focus on what Quebec has in common with similar phenomena in many other countries.

The cross-country study of generic phenomena develops data that are logically and to a considerable extent psychologically equivalent to the background information that we tend to ignore in the lawyer/engineer and Vietnam/Cambodian aircraft experiments. As previously noted, we use such information when it appears causally relevant to the outcome being estimated, and we ignore when it does not. Developments in other feudal societies subjected to modernizing influences do not cause events in Saudi Arabia, just as the activities of ethnic movements in a half dozen or so other western countries do not cause developments in Quebec (except to the extent that the power of successful example causes the spread of certain tactics). Consequently, this background information is not considered significant by analysts accustomed to thinking in terms of immediate causes. Of course, the purpose of cross-national research is to identify common causes at work in a number of comparable situations. But such causes appear to the country analyst immersed in case-specific data to be abstract and distant as compared with the valid, concrete, and seemingly compelling evidence about the specific case at hand.

ANALYSIS ACTIVITIES
Biases in Evaluating Information

ABSENCE OF EVIDENCE

A principal characteristic of intelligence analysis is that key information is generally lacking. Analytical problems are selected on the basis of their importance and the perceived needs of the consumers, without much regard for availability of information. We have to do the best we can with limited information. Although this certainly ought not imply being content with just the reporting that flows through the in-box, we must anticipate large gaps and somehow take into account the fact that much relevant information is known to be missing.

Ideally, intelligence analysts should recognize what relevant evidence is lacking and be able to factor this into their calculations, estimating the potential impact of the missing data and adjusting confidence in their judgment downward in recognition that key information is unavailable. Unfortunately, this ideal may not be the norm. "Out of sight, out of mind" may be a better description of the impact of gaps in the evidence.

This problem can be demonstrated using a fault tree, which is a schematic drawing showing all the things that might go wrong with any endeavor. Fault trees are often used to study the fallibility of complex systems such as a nuclear reactor or space capsule. Figure 49 is a fault tree showing all the reasons why an automobile might not start.(11)

The "car won't start" fault tree was shown to several groups of experienced mechanics. One group was shown the full tree and asked to imagine 100 cases in which a car won't start. Members of this group were then asked to estimate how many of the 100 cases were attributable to each of the seven major branches of the tree, that is, to battery failure, ignition system failures, etc. A second group of mechanics was shown only an incomplete version of the tree; three major branches were omitted in order to test how sensitive the test subjects were to what was left out.

If the mechanics' judgment had been fully sensitive to the missing information, then the number of cases of failure that would normally be attributed to the omitted branches should have been added to the "Other Problems" category. In practice, however, the "Other Problems" category was increased by only half as much as it should have been, indicating that the mechanics shown the incomplete tree were unable to fully recognize and incorporate into their judgments the fact that some of the causes for a car not starting were missing from the fault tree. When the same experiment was run with non-mechanics, the effect of the missing branches was much greater.

As compared with most questions of intelligence analysis. the car won't start experiment involved rather simple analytical judgments concerning information that was presented in a well-organized manner. That the presentation of relevant variables in the abbreviated fault tree was incomplete could and should have been recognized easily by the experienced mechanics selected as test subjects. That the mechanics performed so poorly on this experiment suggests that intelligence analysts may have

FIGURE 49 -- FAULT TREE FOR NON-STARTING CAR PROBLEM

```
                          CAR WON'T START
                                 |
                                 |
BATTERY CHARGE INSUFFICIENT -----+----- IGNITION SYSTEM DEFECTIVE
    Ground connections faulty    |          Coil faulty
    Terminals loose/corroded     |          Distributor faulty
    Battery weak                 |          Spark plugs defective
                                 |          Wiring between components
STARTING SYSTEM DEFECTIVE -------|              defective
    Switches defective           |
    Transmission not in park     |----- OTHER ENGINE PROBLEMS
        or neutral               |          Oil too thick
    Seat belt problem            |          Pistons frozen
        (1974 cars)              |          Compression poor
    Starter motor faulty         |
    Starter drive defective      |----- MISCHIEVOUS ACTS OR VANDALISM
                                 |          Stolen or broken vital parts
FUEL SYSTEM DEFECTIVE -----------|          Siphoning of gas
    Insufficient fuel            |          Disruption of wiring
    Excess fuel (flooding)       |
    Defective choke              |----- ALL OTHER PROBLEMS
    Defective air filter         |
```

similar problems. Missing data is a normal characteristic of intelligence problems, and it is probably more difficult to recognize and incorporate the missing data into judgments on abstract intelligence problems that in the more concrete "car won't start" experiment.

As an antidote for this problem, it would seem important to identify explicitly those relevant variables on which information is lacking, consider alternative hypotheses concerning the true status of these variables, then modify our judgment and especially confidence in our judgment accordingly. It is also necessary to consider whether this lack of information on such matters is normal or whether absence of information is itself an indicator of unusual activity or inactivity.

ANALYSIS ACTIVITIES
Biases in Evaluating Information

OVERSENSITIVITY TO CONSISTENCY

Internal consistency in a pattern of evidence is a major determinant of our confidence in judgments based on that evidence.

For example, "...people express more confidence in predicting the final grade-point average of a student whose first-year record consists entirely of B's than in predicting the grade-point average of a student whose first-year record includes many A's and C's."(12) This confidence in illusory. Similarly, when test subjects were asked to estimate class standing, their estimates were almost identical when based on knowledge that a student received a single B in one course as when based on having an A in one course and a C in another; however, there was significant difference in their level of confidence in the estimate. Subjects expressed far more confidence when predicting from a single grade in one course than from an inconsistent pair of grades.(13) This is not justifiable, as the A and C are equivalent to two B's and the larger sample (two grades versus only one) is clearly a more valid basis for estimation.

In one sense, consistency is clearly an appropriate guideline for evaluating evidence. We formulate alternative explanations or estimates and select the one that encompasses the greatest amount of evidence within a logically consistent scenario. Under some circumstances, however, consistency can be deceptive. Information may be consistent only because it is highly correlated or redundant, in which case many related reports may be no more informative than a single one of them. Or it may be consistent only because we are drawing our information from a very small sample or a biased sample.

Such problems are most likely to arise in intelligence analysis when we have little information, say on political attitudes of Soviet military officers or among certain African tribes. If the available evidence is consistent, we will commonly overlook the fact that it represents a very small and hence unreliable sample taken from a large and heterogeneous group. This is not simply a matter of necessity -- having to work with the information on hand, however imperfect it may be; rather there is an illusion of validity caused by the consistency of the information.

The tendency to place too much reliance on small samples has been dubbed the "law of small numbers."(14) This is a parody on the law of large numbers, which is a basic statistical principle that says large samples will be highly representative of the population from which they are drawn. This is the principle that underlies opinion polling. But it appears that people are not good intuitive statisticians. We do not have much intuitive feel for how large a sample has to be before we can draw valid conclusions from it. The law of small numbers means that, intuitively, we treat small samples as though they were large ones.

This has been shown to be true even with mathematical psychologists with extensive training in statistics. Psychologists designing experiments have seriously incorrect notions about the amount of error and unreliability inherent in small samples of data, unwarranted confidence in the early trends from the first few data points, and unreasonably high expectations of being able to repeat the same experiment and get the same results with a different set of test subjects.

I suspect that as intelligence analysts we make the same kinds of mistakes. We are overly confident of conclusions drawn from very little data -- especially if that data seems to be consistent. When working with a small but consistent body of evidence, analysts need to consider how representative that evidence is of the total body of potentially available information. If more reporting were available, how likely is it that this information, too, would be consistent with the already available evidence? If one is stuck with only a small amount evidence and cannot determine how representative this evidence is, say, of attitudes within the Soviet military, confidence in judgments based on this evidence should be low regardless of the consistency of the information.

COPING WITH LESS-THAN-PERFECTLY-RELIABLE EVIDENCE

There are many reasons why information may be less than perfectly reliable: misperception or bias on the part of the ultimate source; distortion in the reporting chain from subsource through source (case officer, reports officer to the analyst); and misunderstanding or misperception by the analyst. Further, much of the evidence we bring to bear in conducting analysis is retrieved from memory, but we often cannot remember even the source of information we have in memory let alone the degree of reliability we attributed to that information when it was first received.

As mentioned repeatedly in other articles in this series, the human mind has difficulty coping with complicated probabilistic relationships so we tend to employ some simple rules of thumb that reduce the burden of processing such information. In processing less-than-perfectly-reliable information, we do focus on source reliability, but we tend to make a simple yes -- no decision. If we reject the evidence as unreliable, we tend to reject it fully so it plays no further role in our mental calculations. If we accept it, we tend in our subsequent mental calculations to accept it as wholly reliable, ignoring the probabilistic nature of the reliability judgment. This is called a best guess strategy;(15) such a strategy simplifies the integration of probabilistic information, but at the expense of ignoring some of the uncertainty. If we have information about which we are 70 percent or 80 percent sure but treat this information as though it were 100 percent certain, judgments based on that information will be overconfident.

A more sophisticated strategy is to make a judgment based on an assumption that the available evidence is is perfectly reliable, then to reduce the confidence in this judgment by a factor determined by the assessed reliability of the information. For example, the available evidence may indicate that a certain event will probably (75 percent) occur, but we cannot be certain that the evidence on this judgment is based is wholly accurate. Therefore, we reduce the assessed probability of the event (say, down to 60 percent) in order to take into account the uncertainty concerning the evidence. This is an improvement over the best guess strategy but generally still results in judgments that are overconfident when compared with the mathematical formula for calculating probabilities.(16)

In mathematical terms, the joint probability of two events is equal to the product of their individual probabilities. Imagine a situation in which you receive a report that is probably (75 percent) true. If the report is true, you judge that even X will probably (75 percent) happen. The actual probability of event X is only 56 percent, (which is derived by multiplying 75 percent by 75 percent).

In practice, life is not nearly so simple. Analysts generally must consider many items of evidence with different degrees of reliability that are related in complex ways with varying degrees of probability to several potential outcomes. Clearly, one cannot make neat mathematical or even intuitive calculations that take all of these probabilistic relationships into account. In making intuitive judgments, we unconsciously seek shortcuts for sorting through this maze, and these shortcuts involve some degree of ignoring the uncertainty inherent in less-than-perfectly-reliable information. There seems to be little an analyst can do about this short of breaking the analytical problem down in a way that permits assigning probabilities to individual items of information, and then using a mathematical formula to integrate these separate probability judgments.

PERSISTENCE OF IMPRESSIONS BASED ON DISCREDITED EVIDENCE

Impressions tend to persist even after the evidence that created those impressions has been fully discredited. Psychologists have become interested in this phenomenon because many of their experiments require that the test subjects be deceived -- for example, that they be made to believe they were successful or unsuccessful in performing some task or that they possess certain abilities or personality traits when this is not in fact the case. Professional ethics require that test subjects be disabused of these false impressions at the end of the experiment, but this has proven surprisingly difficult to achieve.

Students' erroneous impressions concerning their logical problem-solving abilities persevered even after the students were informed that manipulation of good or poor teaching performance had virtually guaranteed their success or failure.(17) Similarly, test subjects asked to distinguish true from fictitious suicide notes were given feedback that had no relationship to actual performance; the test subjects had been randomly divided into two groups, with members of one group being given the impression of above average success and the other of relative failure at this task. The subject's erroneous impressions of the difficulty of the task and of their own performance persisted even after they were informed of the deception, that is, informed that their alleged performance had been preordained by their assignment to one or the other test group. Moreover, the same phenomenon was found among observers of the experiment as well as the immediate participants.(18) The impressions persisted even after the evidence on which they were based was fully discredited.

There are several cognitive processes that might account for this phenomenon. The tendency to interpret new information in the context of pre-existing impressions is relevant but probably not sufficient to explain why the pre-existing impression cannot be eradicated even when the new information authoritatively discredits the evidence on which it is based. An interesting but speculative explanation draws on the strong human tendency to seek causal explanations.

When evidence is first received, it is perceived within a context that implies causal connections between the evidence and some antecedents that explain the evidence. The stronger the perceived causal linkage between the evidence and its antecedents, the stronger the impression created by the evidence. Thus in the experiment with suicide notes, one test subject attributed her apparent success in distinguishing real from fictitious notes to her empathic personality and the insights she gained from the writings of a novelist who committed suicide. Another ascribed her apparent failure to lack of familiarity with people who might contemplate suicide. Even after learning that the feedback concerning their performance was invalid, these subjects retained this plausible basis for inferring that they were either well or poorly qualified for the task. Thus their initial impressions of task difficulty and of their own ability remained unchanged.(19)

In more general terms, when evidence is received, we postulate a set of causal connections that explains this evidence. Even though the evidence may subsequently be discredited, the causal linkages remain plausible and may be seen as sufficient to imply the existence of an event even in the absence of the now discredited evidence. The previously perceived causal linkage comes easily to mind. It is a readily available(20) explanation that makes the event seem more likely than it would have appeared prior to receipt of the discredited evidence. Colloquially, we might say that once information rings a bell, the bell cannot be "unrung."

The ambiguity of most real world situations contributes to the operation of this perseverance phenomenon. Rarely in the real world is evidence so thoroughly discredited as is possible in the experimental laboratory. Imagine, for example, that you are told that a clandestine source who has been providing information for some time is actually under hostile control. Imagine further that you have formed a number of impressions on the basis of reporting from this source. It is easily to rationalize maintaining these impressions by arguing that the information was true despite the source being under control, or by doubting the validity of the report claiming the source to be under control. In the latter case, the phenomenon of impression perseverance may itself affect evaluation of the evidence that supposedly discredits the impression; it is because we retain our initial impression concerning the validity of the information that we disbelieve the new evidence concerning the source.

This phenomenon also relates to the conduct and detection of deception. Conventional wisdom holds that security is an essential element of deception. Compromise of a deception plan might be worse than no deception at all, for it could attract attention to the true situation. Yet a study of 68 cases of strategic deception and surprise between 1914 and 1968 found significant security leaks in every case, but not a single instance in which the deception failed as a result.(21) One reason for this is the incredible ability humans have to rationalize discrepant information to fit their preconceptions. A second and more surprising reason is this tendency for impressions to persist even though the source of information on which they were based has been discredited.

The process that causes this phenomenon may also affect our reaction to information that is plausible but known from the beginning to be of questionable authenticity. Ostensibly private statements by foreign officials are often reported through intelligence channels with the caveat that the source may have known the information would be passed to the US government. In other words, it is not clear whether a private statement by an ambassador, cabinet member or whomever is an actual statement of private views, an indiscretion, or a deliberate attempt to manipulate the perceptions of US Government officials.

The analyst who receives such a report with a sterilized source description often has little basis for judging the source's motivation, so the information must be judged on its own merits. In making such an assessment, the analyst is influenced by plausible causal linkages. If these are linkages of which the analyst was already aware, the report has little impact as it simply supports existing views. If there are plausible new linkages, however, thinking is restructured to take these into account. It seems likely that the amount of impact on our thinking is determined solely by the substance of the information, and that the caveat concerning the source does not attenuate the impact of the information at all. Knowing that the information comes from an uncontrolled source who may be trying to manipulate us does not necessarily reduce the impact of the information.

REFERENCES

1. Most of the ideas and examples in this section are from Richard Nisbett and Lee Ross, Human Inference: Strategies and Shortcomings of Social Judgment (Englewood Cliffs, N.J.: Prentice-Hall, 1980), Chapter 3.

PROBLEM SOLVING TECHNIQUES
Biases in Evaluating Information

2. A. Paivio, Imagery and Verbal Processes (New York: Holt, Rinehart & Winston, 1971).

3. See "Biases in Estimating Probabilities: The Availability Rule," Analytical Methods Review (March 1980), pp. 14-21.

4. Timothy Crouse, The Boys on the Bus (New York: Ballantine 1974).

5. Nisbett and Ross, op. cit., p. 56.

6. Nisbett and Ross, op. cit., p. 57.

7. Experiment adapted from Daniel Kahneman and Amos Tversky "On Prediction and Judgment," Oregon Research Bulletin, 12,14 (1972).

8. This is a modified version, developed by Frank J. Stech, of the blue and green taxicab question used by Kahneman and Tversky, op. cit.

9. Maya Bar-Hillel, "The Base-Rate Fallacy in Probability Judgments," Acta Psychologica, in press, 1980.

10. See "Strategies for Analytical Judgment," Analytical Methods Review (November 1980), pp. 1-24..

11. Baruch Fischhoff, Paul Slovic and Sarah Lichtenstein, Fault Trees: Sensitivity of Estimated Failure Probabilities to Problem Representation, Technical Report PTR-1042-77-8 (Eugene, Ore.: Decision Research, 1977).

12. Amos Tversky and Daniel Kahneman, "Judgment under Uncertainty: Heuristics and Biases," Science, Vol. 185 (27 September 1974), p. 1126.

13. Daniel Kahneman and Amos Tversky, "Intuitive Prediction: Biases and Corrective Procedures," TIMS Studies in the Management Sciences, 12 (1979), pp. 313-327.

14. Tversky and Kahneman, op. cit., pp. 1125-1126.

15. See Charles F. Gettys, Clinton W. Kelley II and Cameron Peterson, "The Best Guess Hypothesis in Multistage Inference," Organizational Behavior and Human Performance, 10, 3 (1973), pp. 364-373; and David A. Schum and Wesley M. DuCharme, "Comments on the Relationship Between the Impact and the Reliability of Evidence," Organizational Behavior and Human Performance, 6 (1971), pp. 111-131.

16. Edgar M. Johnson, "The Effect of Data Source Reliability on Intuitive Inference," Technical Paper 251 (Arlington, Va.: US Army Research Institute for the Behavioral and Social Sciences, 1974).

17. R. R. Lau, M. R. Lepper, and L. Ross, "Persistence of Inaccurate and Discredited Personal Impressions: A Field Demonstration of Attributional Perseverance," paper presented at 56th Annual Meeting of the Western Psychological Association (Los Angeles, April 1976).

18. Lee Ross, Mark R. Lepper, Michael Hubbard, "Perseverance in Self-Perception and Social Perception: Biased Attributional Processes in the Debriefing Paradigm," Journal of Personality and Social Psychology, 32, 5 (1975), pp. 880-892.

19. Lee Ross, Mark R. Lepper, Fritz Strack and Julia Steinmetz, "Social Explanation and Social Expectation: Effects of Real and Hypothetical Explanations on Subjective Likelihood," Journal of Personality and Social Psychology, 33, 11 (1977), p. 818.

20. For discussion of availability bias, see "Biases in Estimating Probabilities: The Availability Rule," Analytical Methods Review, op. cit.

21. Barton Whaley, Strategem, Deception and Surprise in War (MIT, unpublished manuscript, 1969), p. 164.

PROBLEM SOLVING TECHNIQUES

Biases in Evaluating Information

USES OF FORMAL LOGIC IN INTELLIGENCE

It can be argued that two basic types of logic enter into the business of intelligence analysis. On the one hand there is a kind of informal logic, the sum and substance of which is the creation of a product that is viewed by the majority of consumers as persuasive, as "logical" and not faulty in its reasoning. We all strive for this.

The finished intelligence product also should be based on the principles of formal logic, a topic to which we seldom give much thought. If the product diverges strongly from the tenets of formal logic, it will lose credibility in the eyes of the consumer, regardless of the consumers depth of knowledge in the subject matter of the product or the intracacies of formal logic. It therefore seems useful for analysts to be familiar with the principles of formal logic, both as a tool to "tighten" the credibility of their publications, and (more importantly) as a tool to improve the clarity, accuracy, and justification of their analytical arguments.

It is wise to remember that there is a difference between the finished product and the process one goes through to arrive at the product. Formal logic is useful in evaluating a finished product, and is useful in reviewing the logical consistency of an analytical argument; but the analytical process itself (particularly if one is using heuristic or creative techniques) can be very amorphous and subject to all of the tricks (rules of thumb) and intuition that an analyst can use to make correct and meaningful judgments. It is when the analyst desires to use judgements and arguments in a finished product that careful attention to formal logic should come into play. This is not to say that one's analysis should be illogical, or that logically attacking a problem is bad or inefficient; rather, it is saying that one need not constrain ones thought process with the details of formal logic (or any other detailed formalism). Instead, knowing that these constraints can be used later to ensure consistency and to measure credibility, one can proceed full speed ahead and worry about details later.

PROBLEM SOLVING TECHNIQUES
Formal Logic in Intelligence

LOGICAL FALLICIES

A fallacy is any logical defective argument that is capable of misleading people into thinking that it is logically correct.**

FALLACIES OF RELEVANCE

A fallacy of relevance is any logically defective argument in which the premises are unrelated or not pertinent to--and thus incapable of establishing the truth of--its conclusion.

Argument by Force

A person appeals to force or to the threat of force--as contrasted to offering reason--to cause the acceptance of a conclusion.

Ad Hominem (Abusive Variant)

A person directs an argument against the truth of what another person has asserted by attacking the person who made the assertion, rather than attacking the merits of the assertion itself.

** The Logical Fallicies section has been adapted from material developed by Mr. David E. Farnham for use in his Arlington Hall Station Logic for Intelligence Analysts course.

PRODUCTION ACTIVITIES
Formal Logic in Intelligence

Ad Hominem (Circumstantial Variant)

One disputant or contender ignores the question of whether his own contention is true or false; and, instead, seeks to prove that his view ought to be accepted because of the other contender's special circumstances (such as culture, religion, occupation, technical expertise, politics, etc.).

Ad Hominem ("You Too" Variant)

Trying to ridicule, humiliate, or embarrass a person during a debate or some other adversary forum by claiming inconsistency or special pleasing rather than by repudiating his position with reasons.

Argument from Ignorance

An argument is claimed to be valid (or a proposition true) simply on the basis that it has not been proved invalid (or the proposition false).

Argument for Pity

Sympathy is appealed to, rather than reasons given, in order to gain acceptance for a conclusion.

Argument to the People

Attempting to win popular assent for a conclusion by arousing the passions or enthusiasm of the people, multitudes, or crowds.

Argument from Authority

Appealing to some authority for testimony in matters outside the province of his/her special field of expertise.

Accident

Attempting to apply a general rule to a particular case whose special or accidental circumstances render the rule inapplicable.

PROBLEM SOLVING TECHNIQUES

Formal Logic in Intelligence

Converse Accident

A general conclusion is drawn on the basis or unusual or atypical cases.

False Cause (Variant #1)

Identifying the wrong cause for a given event.

False Cause (Variant #2)

Identifying the cause of an event to be the true cause simply on the basis that it occurs prior to the event.

Begging the Question

One assumes as a premis in an argument the very conclusion intended to be proved.

Complex Question

Seeking a single conclusion from a question that has two or more distinct parts.

Irrelevant Conclusion

An argument purported to establish a particular conclusion is directed to proving a different conclusion.

Statistical

Misuse is made of statistical methods to prove a conclusion.

PRODUCTION ACTIVITIES

Formal Logic in Intelligence

FALLACIES OF AMBIGUITY

A fallacy of ambiguity is any logically defective argument which contains words or phrases whose meaning shifts or changes, more or less subtly, in the course of formulation and thus renders it fallacious.

Equivocation

A different meaning of a single word or phrase is introduced in the development of an argument with the result that such words or phrases have two or more meanings.

Amphiboly

Arguing from premises that have ambiguous formulations because of the grammatical construction used or because of the loose and awkward way in which words are combined.

Accent

The meaning of a proposition is changed due to emphasizing, stressing, or placing accent on some particular word or phrase in the proposition.

Division

When reasoning incorrectly from that which is an attribute:

* of the whole to that which is an attribute of a part of the whole.
* of the whole collectively to that which is an attribute of each element or member of that collection.

PROBLEM SOLVING TECHNIQUES

Formal Logic in Intelligence

Composition

When reasoning incorrectly from attributes of:

* parts of the whole to attributes of the whole itself.
* the individual elements of a collection to the attributes of the total elements or members of that collection.

FALLACIES OF ANALOGY

An Analogical Fallacy occurs whenever it is incorrectly inferred that if two or more entities are similar to one another in one or more respects, then they are probably similar in other respects as well.

Perfect Analogy

An analogical fallacy which occurs when reasoning from a partial resemblance between two entities to an entire and exact correspondence between those same entities.

Prediction by Analogy

A fallacy that occurs when analogy is used to anticipate future events in which there is an indeterminate probability that renders prediction utterly untestable and inconclusive.

False Analogy

An analogical fallacy which occurs when the terms are shifted from one analog to another.

EFFECTIVE WRITING AND BRIEFING TECHNIQUES

GENERAL

Briefings and written publications usually have one main purpose. This purpose should be clear in your mind when you prepare your product, and also should be clear to the audience in the product you deliver. Typical purposes include:

a. Providing Information

b. Teaching or Training

c. Convincing

d. Impacting a Decision

Because intelligence products provide information, our audiences do expect to hear clear, concisely stated conclusions and projections. They also expect logical and objective arguments. Detailed facts may or may not be appropriate, depending on the scope of the topic and the technical sophistication, interest, and need of the audience.

But to be effective, each product must have, in addition to a clear purpose, a technical content and organization specifically tailored to its audience. This involves using a vocabulary and thought process familiar to your audience. For most of us, this requires developing new communication skills.

The formal communication (writing and briefing) skills most of us learned in school usually involved presenting material to a teacher. Our main purpose was impressing upon someone who knew more than we did about something, how well we too had mastered the subject area. The audience in a work environment changes drastically from that of a school environment. Consequently, the style and technical content of our products should chance too.

At work, the expert is the <u>author</u>, not the <u>reader</u>. Hence, we must write to inform rather than to impress. This means that we must make certain that our intended message is absolutely clear <u>to our audience</u>. We cannot assume that the reader will be familiar with technical jargon, nor with the consequences or implications of observations or calculations. To communicate effectively, we must use vocabulary familiar to both ourselves and the reader. We can introduce new terms, but we must start from some common ground, some common understanding. It is this "commonness" of understanding that permits us to introduce something new. This "new" then becomes common and can be used in turn to introduce something else new. We must work from "old" to "new," and we must do so in a logical, easy to follow manner.

PROBLEM SOLVING TECHNIQUES
Effective Writing and Briefing Techniques

Good writing[50] is not easy. It is hard work. But hard work on our part is necessary to make it easy for the reader to digest the message that is intended. If you as the writer care whether the reader gets your message, it is a dangerous gamble to make the reader work too. The reader always has the choice of not reading anything at all. It is essential, therefore, that we make a reader's work as easy and pleasant as possible. If a paper is work to read, chances are it won't get read.

Statistics show that most readers will look at the summary of a paper and that a slight majority will read the preliminaries, but few read the technical discussion or the appendices. Since, as writers, we spend a great deal of effort on the body of the report, why not raise the odds of getting this portion of a paper read. The only good ways I know to do this, given that one has no authority over a reader, are to make your prose fun to read and your messages obvious.

To help you improve[51] the effectiveness of your formal communication skills, the remainder of this chapter provides some hints toward this end.

EFFECTIVE WRITING TECHNIQUES

THE GUNNING FOG INDEX

Robert Gunning's Fog Index measures the readability of writing. It does this by indexing a writing sample according to the number of years of formal education an average reader would need to read the indexed sample with ease and understanding.

In general, the Fog Index takes two factors into account: (1) the length of sentences (2) the complexity of the words used.

Although not infallible, it is a good tool for assessing how well words were picked and how clearly sentences were constructed. The Fog Index may be applied to any kind of writing except poetry.

To find the Fog Index of a passage, follow these steps:

---

[50] This discussion applies both to written text and to oral briefings, but to improve readability only written products are mentioned explicitly.

[51] If you ever want to feel paranoid, try writing a chapter on effective writing techniques. Fortunately, I am neither required nor famed to be a good writer. Hence, as long as there is some content to this chapter I will survive the knowledge that it probably is written poorly.

PRODUCTION ACTIVITIES
Effective Writing and Briefing Techniques

One: Take a sample of at least 100 words that begins and ends with a complete sentence. Determine the average number of words per sentence. (Treat independent causes as separate sentences.)

Two: Determine the number of words with three syllables or more per 100 words. Do not count proper nouns, combinations of short easy words like "manpower" and "insofar," and verbs made into three syllables by adding "es", "ed", or "ing".

Add together the answers to steps one and two, and multiple by 0.4. THE RESULT WILL BE THE YEARS OF FORMAL EDUCATION A READER OF AVERAGE INTELLIGENCE WOULD NEED TO READ THAT SAMPLE WITH EASE AND UNDERSTANDING.

Remember that the Fog Index measures the difficulty of reading what is written. It does .us not measure the ideas the writing expresses. Thus, a Fog Index of 9 does not mean you are writing at a ninth grade level. Indeed, studies already show that the more difficult the subject, the greater the need for simple writing. This allows readers to concentrate on your .us ideas, not your .us words.

The message of the Fog Index is simple. Choose simple words over hard ones, and keep sentences short. Where do you start? If you are a typical writer, you will find that long words do more to raise your Fog index than do the number of words per sentence. So start questioning polysyllabic words. Maybe you really do need them. But make sure you do before you use them. At the same time, break up long sentences into sets of simple thoughts.

Most popular writing for adults should have a Fog index of 9 or 10; business writing about 11 or 12. (The U.S. average for business writing is 16.)

SOME TYPICAL FI's

True Confessions - 7	Wall Street Journal - 10
Ladies Home Journal - 8	Harper's - 11
Reader's Digest - 9	Atlantic Monthly - 12

CHECKLIST FOR REVIEWING IN-DEPTH PUBLICATIONS

This checklist is adapted from one developed by Diana Tansill of the OSWR Publications Group for use by office branch chiefs. It assumes that a review addresses substantive and technical accuracy, as well as writing style.

Front Section (FS = Key Judgments or Summary)
Is the purpose of the paper clear to the "unwitting reader?" Does the FS contain implications--where this paper fits in the big picture--of what it means to the US Government, specifically national security?

PROBLEM SOLVING TECHNIQUES
Effective Writing and Briefing Techniques

Are the conclusions the appropriate ones? Are they consistent with the text? Are any significant ones left buried in the text? Are there inappropriate historical inclusions/conclusions in the FS? Are the conclusions supported in the text?

Do the conclusions contain specific evaluations and explanations sufficient for the lay reader? Do you tell the reader not only what the specific frequency is that you have found for the radar but also what this specific frequency means in terms of mission or role?

Are projections (future developments) addressed in the FS? Is the FS brave, i.e., does it lean forward?

Do you have a balance of evidence, judgment, and implications in the FS?

Is the FS short (no more than about 10 percent of the length of the basic text) and is it tightly written--directly to the point?

Boby

Are generic sources sufficiently cited so that it is clear what is fact and what is author's judgment? (Samples of generic source statements are: "special intelligence shows," "satellite imagery indicates," "a reliable source recently stated," and "previously published assessments revealed.")

Are judgmental statements in the body quickly followed up with the evidence and the rationale that led the author to make the judgment? (Don't make the reader wonder "How do you know?".) Are the assumptions specifically stated?

Is the body so arranged that a reader can look at the Contents page and easily know where to look for the backup for the judgments contained in the FS?

Has the author made sure that the paper does not end in a whimper? Are the gaps, future projections, and implications briefly summarized at the end of the paper?

Has the author checked for "mechanical" correctness? Are the paragraphs classified? Is the source listing complete? Are references keyed to the text?

Is the paper sufficiently short that it will get read? (From 8-40 pages of double-spaced draft have been suggested as an appropriate length for the body of a paper.)

Are appendixes used to present historical, tutorial, or highly technical supporting material (methodology) rather than the main text?

Is extraneous material excluded, despite author's pleas that this is the first time that anyone has pulled this all together? Is everything germane to the intelligence questions being addressed? (If in your judgment the extraneous material will be of lasting value to the analytical community, have you considered a separate Technical Intelligence Report, Reference Aid, or typescript as a means for publishing it?)

Are the lengths of the subsections consonant with their importance?

Does the organization minimize repetition? Are facts, evaluations, and explanation all intermingled appropriately? Is the paper so organized that there is not a major section on evidence (especially photographic intelligence-type descriptions) and then a section on analysis and then one on implications? (Once in a blue moon, this is OK.)

Is the organization logical--like subjects grouped together? Is everything on submarine speed or uranium mining in one place? (If you are stuck on flow, try noting the subject of each paragraph in the margin to see if like subjects are grouped togethor and to see if the subject of the paragraph is clear.)

Does the body emphasize what is being done, how good is it, and why is it being done, rather than who is doing it and where is it being done? (That is, be judgmental and interpretive, not merely descriptive.)

Does the paper contain the latest evidence? Is the cutoff date correct?

Has the author hedged on some questions rather than stating "We don't know?" Are the unknowns and the assumptions clearly identified?

Has redundancy between tables and text been eliminated?

Graphics

Are all the graphics necessary? Have they been thought through to the same degree that the text has? Do they contribute to the reader's understanding of the issue (rather than confuse him)? Have you considered whether they are overly technical and/or complicated?

General

Is the subject of the paper still of importance and interest? Is the evidence too weak or too scanty to continue with processing the paper?

Has the author, at a minimum, discussed the paper with others having an equity or expertise in the subject or, at best, had them review the draft? Have major contributions from others been acknowledged?

Have you considered asking your editors for help, especially in the early stages of the paper?

Does it follow any previously approved outlines?

PROBLEM SOLVING TECHNIQUES
Effective Writing and Briefing Techniques

Have you considered your audience? For whom did you write the paper and what information/conclusions do you expect to provide? Do you think that you succeeded?

Have you read the paper word-for-word from beginning to end?

Is the technical vocabulary appropriate for the anticipated audience? Are esoteric technical words defined? Is simple versus pretentious phraseology used? Are acronyms kept to a minimum?

Have you tried to look for what is not said? Are there areas that should have been covered but were not?

EFFECTIVE BRIEFING TECHNIQUES

WHY ARE YOU GIVING THE BRIEFING?

Briefings usually have one main purpose. Keep this purpose clearly in mind while you prepare and while you present your briefing. Also, match the level of technical content and organization of your briefing, and match your delivery technique to the technical sophistication and interest of your audience. Take the time to find out who will be in your audience so that you can do this effectively.

WHAT CAN YOU DO TO PREPARE?

Know Your Subject

Keep in mind that as a professional, you are expected to be an expert on your subject. It is perfectly acceptable for other experts to exist also, and for you to have professional disagreements with them, so be open and candid about such disagreements. Openness of this kind will help your audience. In contrast, it is not acceptable for you, as an expert, to hide differences of opinion, and it is not acceptable for you to be ignorant of them. You should be aware of what is going on in your area(s) of responsibility. Your audience will (should) feel abused if you pretend to know or be something that you are not.

Know Your Audience

Learn as much about your audience ahead of time as possible. Its size, technical background, interests, decision needs, customs, expected level of formality, and other such things will affect how well your briefing style will be received. Prepare a briefing content and delivery style that will add to, rather than detract from your briefing in the minds of the audience. That is, as much as possible, do it the audience's way. Don't ask or expect the audience to adapt or accept your preferences. If you want them to listen and to understand (and to care?), do everything you can to make things easy and comfortable for your audience.

Outline Your Briefing

It is almost always wise to prepare an outline of the main points you want to cover. This will help you review the content of your talk when preparing, and can help you make sure that you have covered all of the essential points while talking. It also can help you get back on track if something (a question) disrupts your train of thought while speaking.

Do not, however, write out your whole briefing. Simply outline the main points and hi-lite any points that are particularly important. If someone asks a question, particularly if it covers an issue that you planned to address later, you will find it quite difficult to return smoothly to a complete script of a briefing and/or to revise it impromptu to avoid being redundant with questions you've just answered, or to include new material that a question asks you to address. Furthermore, if your briefing is completely written out, you will have a tendence to read it. Reading a briefi is seldom as effective as a more spontaneous delivery.

Plan Your Tactics

Don't be afraid to be sneaky and innovative. If you need more time than allocated for the briefing, schedule an appointment with the key members of your audience for the hour immediately following the briefing. If they get up to leave, you can go with them, or simply tell them not to rush since their next appointment is with you. If you want to address an issue that doesn't quite fit into the briefing topic, or that would be "politically" inappropriate for you to raise, arrange some planted questions in the audience.

Prepare Effective Audio/Visual Aids

Human beings are basically visual animals. Supplement your verbal message with graphical aids or filmstrips or video tapes, etc. But, keep in mind that these aids are to help make the main points of your talk more clear and vivid. Therefore, choose them carefully, and don't overload your audience with viewgraphs and other pictures or audio effects that overload or bore them.

WHAT SHOULD YOU DO WHILE BRIEFING?

Dress Appropriately

PROBLEM SOLVING TECHNIQUES
Effective Writing and Briefing Techniques

Dress in moderate-tone solids or other non-distracting clothing. Bright red suits, loud stripped sport coats, fancy jewelry, the latest fashion, a favorite bright pink tie, etc. may be particularly appealing to you, but will generally be distracting to the audience. Don't wear things that will give them any extra opportunity to get distracted from what you are saying.

Also, dress to the expectations of the audience. If you are giving a formal briefing to an audience wearing three piece suits, don't show up in blue jeans and a tee shirt. You won't win any points with the audience by such a move, and you could very well antagonize them to the point that they won't be willing to listen to you. Why cause extra problems you don't need?

Maintain Good Stage Presence

Be enthusiastic, maintain good posture, vary loudness and pitch of your voice, and match the pace of your delivery to the size of the room and the audience. The human voice usually can vary its pitch through a full octave on a musical scale. It also can speak in continuous gradations of loudness from whispers through loud yelling. Use these tools to give some emotion to your delivery, to highlight key points, and to help the audience stay alert. Also, remember to speak more loudly and more slowly in larger rooms and to larger audiences.

If you have a podium, use it to hold you briefing notes, not your body. Whereas it is acceptable to lean gently against a podium, it should not become a crutch, and you should not be in danger of falling down if someone were suddenly to move it.

Stand (sit) erect and look alert. Try to look healthy and enthusiastic, not frozen and rigid, nor like you are standing a attention in a military formation. If you look good, you'll help your audience feel good.

Be a careful, however, not to get too aggressive in ways that could offend the audience. If you get an antagonistic and hostile question, answer as politely and well as you can an continue on with your talk. The audience will tend to sympathize with the speaker in cases like this. If, on the other hand, you decide to take the questioner on in a mutually antagonistic manner, and especially if you attend to belittle the questioner or to put him down unnecessarily, the audience will usually side with their comrade. Keep them on your side. If the questioner gets to out of hand, the audience will usually take care of the problem for you.

Use Gestures Effectively

Be and act natural. Smile and be cheerful. Use normal gestures and don't be afraid to move around somewhat. It is unlikely that you stand frozen in place when you talk with your friends about topics that interest you. Effective speakers use the same casual mannerisms that we all use in everyday conversation. Don't stop using them and turn into a frozen statue just because you're talking to a group.

GRAPHICAL AND VISUAL AIDS

TIPS FOR MAKING SLIDES OR VIEWGRAPHS

> This section was adapted from the slide preparation instructions given to speakers scheduled to present papers at the June, 1981, Institute of Technologists convention in Atlanta, Georgia.

Although slides are a basic element of technical presentations, they can ruin a presentation because of any one or all of the following reasons:

1. Slides are not legible to the entire audience.
2. Slides are too detailed or complex to read, not to mention
3. Slides are presented on faulty or improper equipment or are not loaded properly.

The guidelines and "do's and dont's" on the following pages cannot cover every aspect for preparing slides but offer enough of the main points to enable anyone to prepare good, if not excellent slides. It is a speaker's responsibility to consider the preparation of slides as important as the preparation of the oral portion of a presentation. Quality slides are a requisite of a quality presentation but the speaker who must apologize for his slides has lost his audience and, perhaps, its respect for the work being presented.

PLAN SLIDES EARLY

A. Slides are basic to a technical presentation. They must:

1. Save time
2. Increase interest
3. Generate and hold attention
4. Clarify an idea
5. Reinforce or emphasize an idea
6. Add an overtone or undertone (such as enlivening basically dull material)
7. Prove a point
8. Increase Memory Retention

PROBLEM SOLVING TECHNIQUES

Graphical Aids

KEEP SLIDES SIMPLE

A. Don't overcrowd slide-the less copy, the better the slide.

   1. Use Keywords only

B. One primary idea per slide.

C. Breakdown the complex slide into several simple ones. Ask yourself these questions:

   1. Will it be visible? -The more detail, the smaller it must be, and the harder it is to see it.
   2. How much will the audience retain?-What is there to retain a large amount of detail in the audience's mind after it leaves the screen.
   3. Is there enough time?-How much of the slide can be read-or will be read-while it is on the screen?

D. When Detail is a must

   1. Break up information into clearly defined categories with as few headings as possible.
   2. Use contrasting colors
   3. Allow plenty of blank space and wide margins
   4. Try covering parts of the information and progressively exposing it in several slides.
   5. Keep slide on screen longer if very complex.
   6. Show the complex slide first, then summarize the detail in as many rudimentary slides as necessary to show over-all concept.
   7. Consider handout material in conjunction with complex slides.

PREPARING ART WORK

A. Hand-drawn slides

   1. Work in Horizontal area with 2:3 ratio.
   2. Use a Leroy lettering pen or equivalent.
   3. Use India ink.
   4. Use a lettering stencil.
   5. Make lines at least 1/16 (.055) inch wide for all charts and graphs.

B. Type-written slides

   1. Art work rectangle should be 3 x 4 1/2 inches.
   2. Make sure all type is brushed clean.
   3. Use an electric typewriter with a carbon ribbon.
   4. Use a piece of carbon paper in reverse so that it leaves an impression on the back side of your art-work.
   5. Use all Upper Case letters.

PRODUCTION ACTIVITIES

6. Do not exceed six (6) double-spaced lines of type or 45 characters and spaces per line.

C. Hand-drawn/wax presstype

1. Art work rectangle 6 x 9 inches.
2. Letters and characters should be at least 1/20 the height of art work-1/16 the height is better. (Art work 6 inches in height requires letters at least .3 inches but .375 inches is better.
3. Make all lines for charts and graphs at least .055 inches wide.

D. Create Contrast

1. Use yellow or pink tinted paper with black letters for positive slides or white paper with black letters for Negative or Colored slides.
2. &LPNegative or &LPColored slides provide best contrast and are comfortable to view.

PHOTOGRAPHING AND PROCESSING

A. Use only Color Slide film.
B. Fill the viewfinder with the art work rectangle.
   1. Be sure to leave a margin.
C. Count slides in "thin glass" or the new thin plastic mounts.
D. Allow Plenty of time for processing.
   1. Have duplicate slides made.
E. If you can read the print on finished 2"x2" slides without a magnifier, then people in rear seats probably will be able to read them when projected on a screen.

PRACTICE

A. Code your script so that you know exact slide location.
B. Practice your presentation with your slides.
C. Practice your presentation . . . .

** On negative (light letters on black background) or colored (colored letters on darker colored background) slides, an optical phenomenon, "Halation," makes letters and characters seem thicker than they really are. Hence, when preparing art work for Negative or Colored slides, letters and characters should be thinner in proportion to their size. Simply, do not use a bold face type for negative slides. On positive slides (light background with dark letters), the opposite is true and a bold type face is much more legible.

D. If you are unable to practice before the meeting, make arrangements to use the IFT's Speaker Ready Room at the meeting.

SLIDE DO'S AND DONT'S

DON'T try to tell the whole story on one slide.

DO express your idea in a few words-like a grade school flashcard or a telegram.

DON'T use heavily complex or detailed slides.

DO use 2,3,5 slides to simplify and communicate the complex idea.

GRAPHICAL DISPLAYS

This section will review briefly several common types of displays that belong to what most of us put in the category of "graphs." They include the standard two axis bar charts and line graphs, as well as pie charts and pictograms.

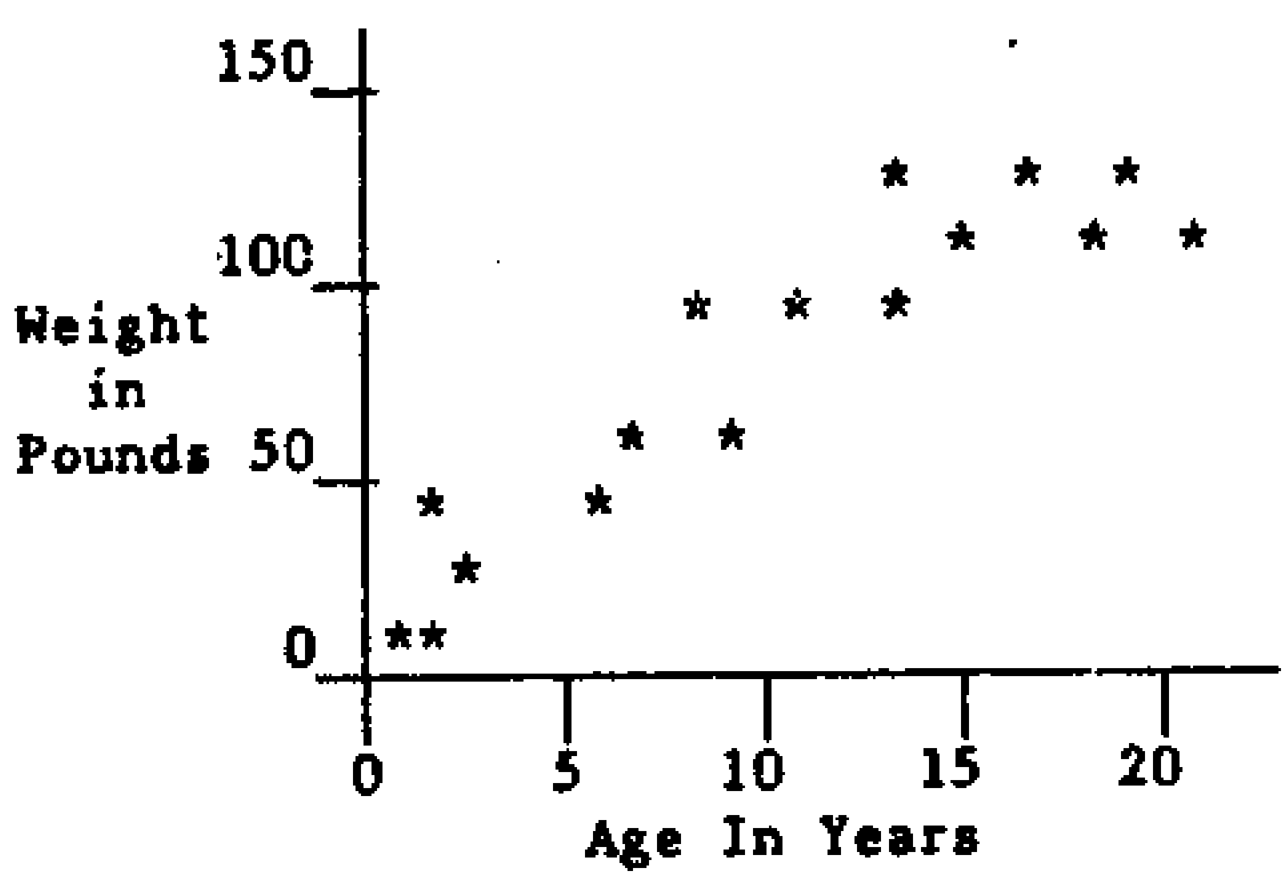
Here in the coursebook we'll basically just define (describe) some of the more common graphical display options since computer graphics on word processing systems is basically nonexistent. We'll leave some blank lines in the text at strategic places for you to copy examples from the lectures, if you so desire.

SCATTER DIAGRAMS

Scatter Diagrams are graphs (usually two, but occasionally three, dimensional) that indicate the value of each and every data point. This is done by using a pair (for the two-dimensional case) of orthogonal axes to indicate the positional value of data points along each axis. The dependent variable is usually displayed along the vertical axis, and the independent variable along the horizontal axis.

Here's a crude example:

FIGURE 50 -- EXAMPLE OF A SCATTER DIAGRAM

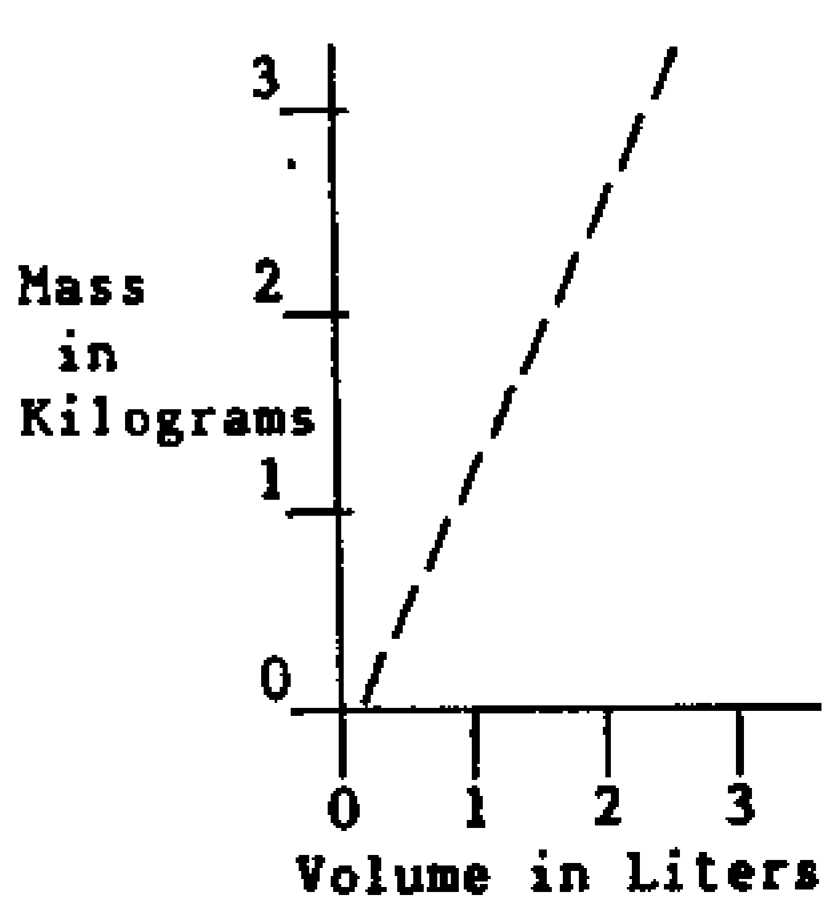


LINE GRAPHS

Line graphs use the same axis structure as scatter diagrams, but some a general trend line rather than indicating the exact value of each data point.

For example:

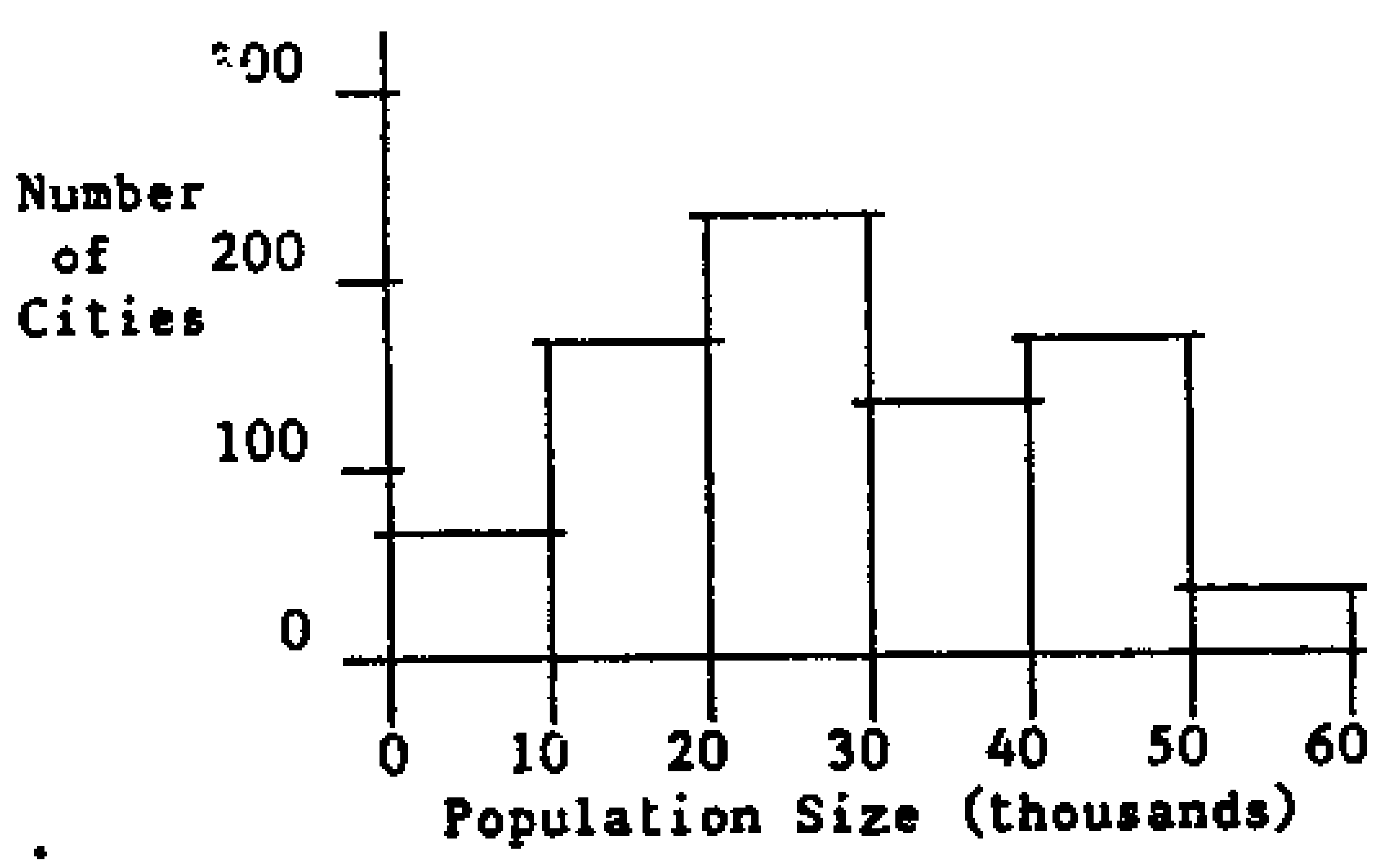
FIGURE 51 -- EXAMPLE LINE GRAPH



HISTOGRAMS

Histograms indicate how many observations (data points) were obtained in each of several contiguous intervals of some measurement scale. The two axis approach is typically used, with the number of observations shown on the vertical axis, and the measurement intervals on the horizontal axis.

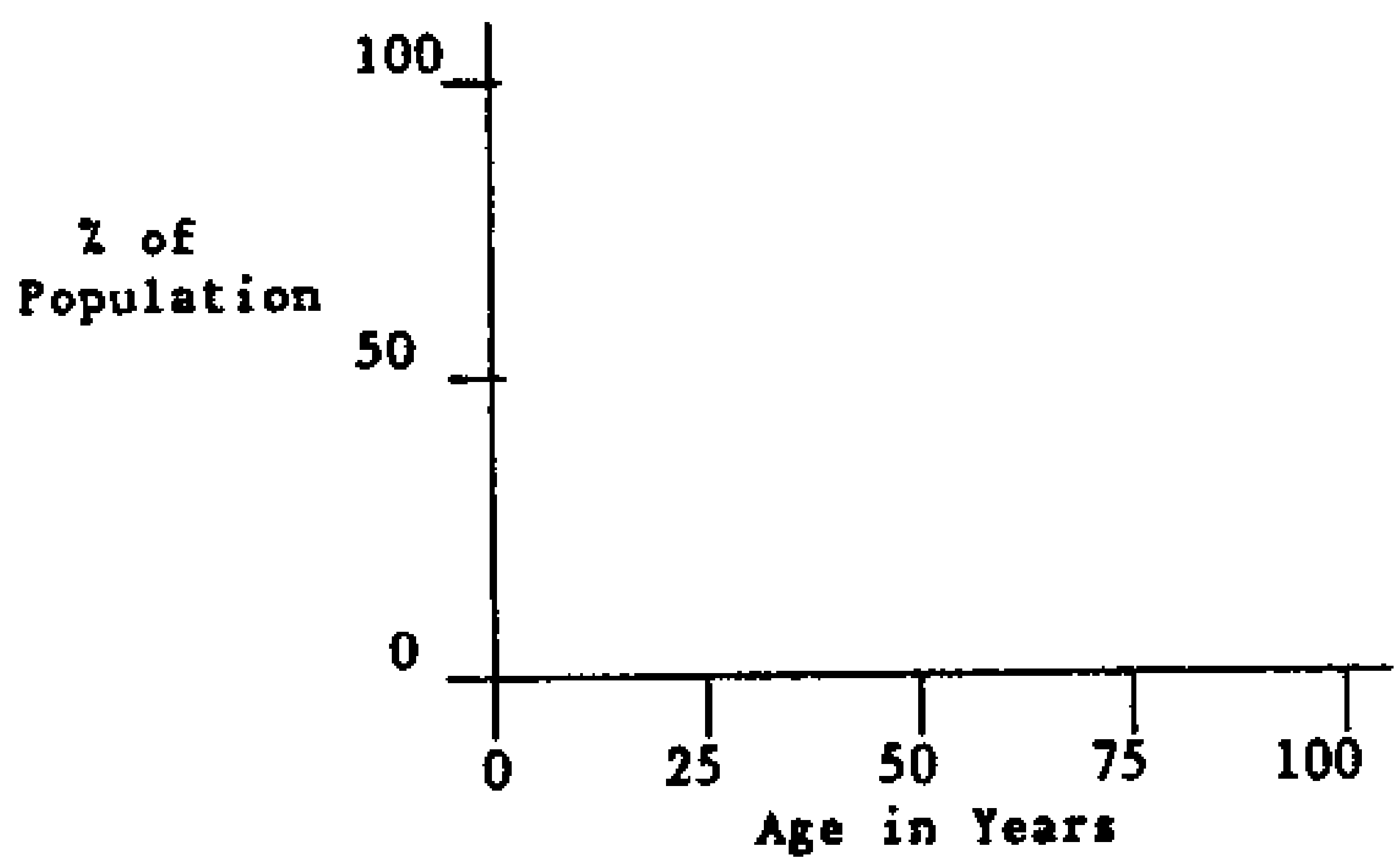
FIGURE 52 -- EXAMPLE HISTOGRAM



CUMULATIVE FREQUENCY DISTRIBUTIONS

Cumulative frequency distributions are line graphs that indicate what fraction (indicated on the vertical axis) of a population (of data) has values less than or equal to a given value on the horizontal axis.

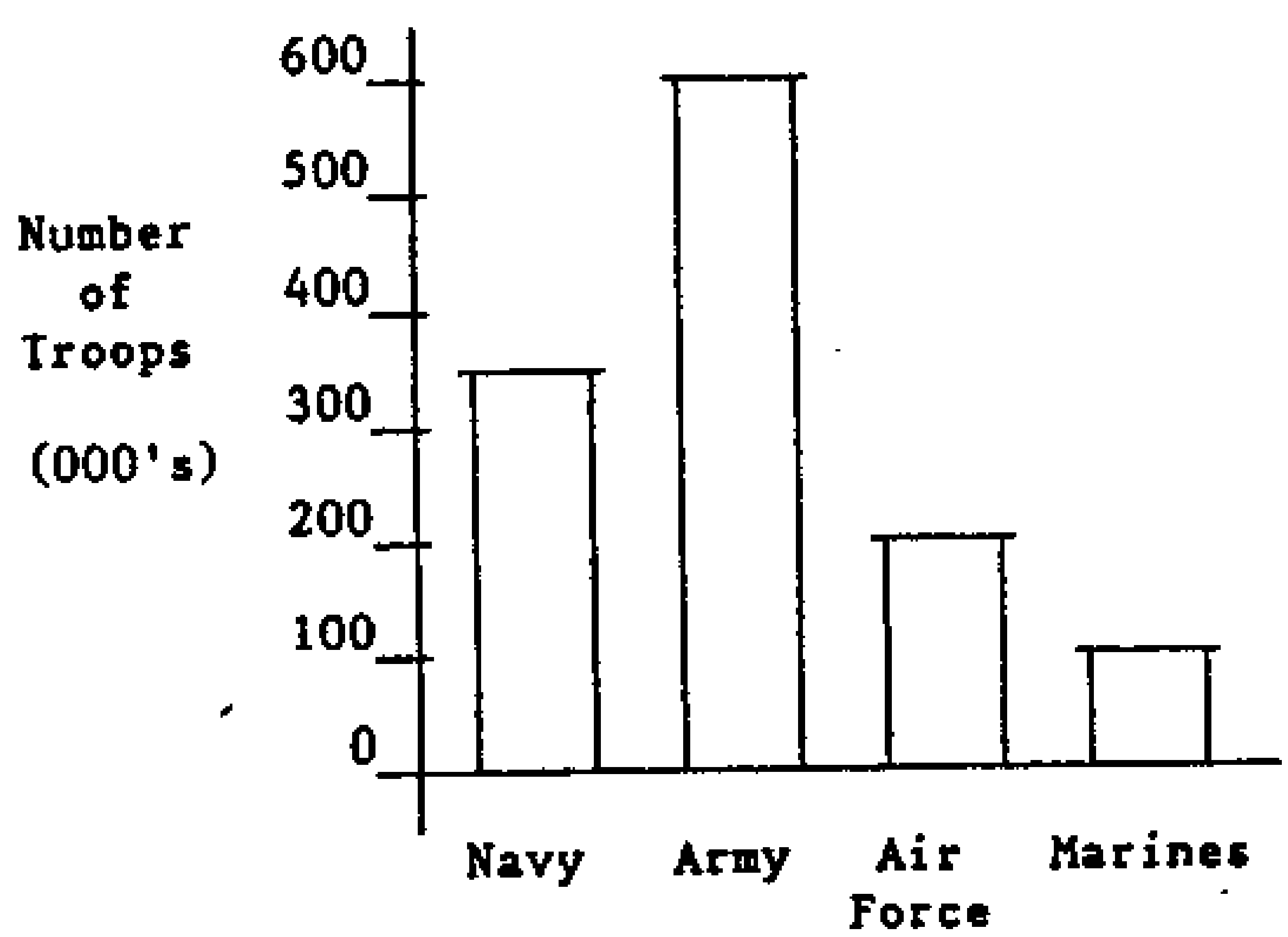
FIGURE 53 -- EXAMPLE CUMULATIVE FREQUENCY DISTRIBUTION



BAR CHARTS

Bar Charts use a histogram-like structure to show the relative size of differing categories of things. Note that, unlike histograms, the "X" (horizontal) axis in a bar chart lists nominal categories. It does not represent a continuous or numerically ordered scale.

FIGURE 54 -- EXAMPLE BAR CHART



PIE CHARTS

Pie charts use a large circle, divided into segments by extending lines from its center to the circumference, to indicate the relative fraction of some population (of data) that is comprised by some particular type of observation. Its use is similar to that of a bar chart.

VENN DIAGRAMS

Venn diagrams are similar to a pie chart in that areas of some geometric figure (often a circle) are enclosed by subboundary lines to indicate what fraction of the total population is comprised by each possible observation (data) type. The major distinctions are that Venn diagrams permit overlapping of observation types (that is, a given observation may belong to multiple observation types) and the subregions can be any shape desired (rather than drawn from the middle with straight lines. This type of display is useful for showing the overlap between different factors (e.g., religion and political preference).

PROBLEM SOLVING TECHNIQUES

PICTOGRAMS

Pictograms are similar to bar charts, except they use a figurine to indicate the relative size of the category observations instead of using a scaled axis. Take care to always use the same sized figurine within each pictogram. Stack (or line up) multiple copies of the same figurine, don't make them bigger or smaller to indicate changes, since your intended meaning will become ambiguous. (Most people are affected by figurine area, which changes as the square of figurine height or diameter.)

GUIDELINES FOR MAKING GRAPHS

1. Graphs should be fully self-explanatory. All information needed to interpret a graph's content should be included in the title, scale labels, and (if needed) footnote(s).

2. The units of measurement and range of measurements on each axis should be indicated clearly. (Traditionally, the horizontal scale--the abscissa--is used for independent variables and the vertical scale--the ordinate--for dependent variables.)

3. Keep graphs simple. Graphs are intended to provide a conceptual overview, not precise numbers. Do not make them confusing by placing numbers in the body of the graph or by trying to show too many curves--a general rule of thumb is to use 3 or fewer curves.

4. Clearly indicate any breaks (or other oddities) in measurement scales.

5. Scale the axes so that the plotted curves give a realistic impression of the trends in the data. A general rule of thumb is to let the height of the vertical scale be about 75 percent the length of the horizontal scale. Also, generally start both scales at zero at a common origin.

6. In the special case of pictograms, remember that viewers generally tend to be influenced by the area of a figure, not just its height. Use two of the base figurines to depict a doubled value, not a figurine twice the height (and width) of the base-size figurine.

TABULAR REPRESENTATIONS

Tabular displays are basically sets of data that are arranged in logical and/or organized structures. Numerous types of common tables formats exist. The following provides some examples.

Having collected or been given a set of data, one must organize it in a way that will display effectively the information contained in the data. The value of presenting the observed variables in the form of well-constructed tables and charts can seldom be overemphasized.

Although this discussion of the construction of tables will by no means be exhaustive, we shall point out the main considerations for tabular presentation and shall establish the vocabulary commonly used.

UNIVARIATE NOMINAL SCALE TABLES

Tables for univariate data with nominal scale characteristics are generally easy to construct, because the data will already have been classified into distinct categories in the collecting process. (By univariate we mean that only one measurement is taken for each observation.) The only effort usually required is counting the number of observations in each category and arranging the results in tabular form. The number of observations occurring in a given category is called the frequency for that category. Table 20 is a typical example.

TABLE 20 -- UNIVARIATE TABLE

Educational Background of 250 Government Leaders

Major Area	Frequency
Arts and Humanities	65
Social Sciences	77
Natural Sciences	32
Business Administration	28
Education	35
Engineering	13

TABLE 21 -- UNIVARIATE TABLE WITH PERCENTAGES

Educational Background of 250 Government Leaders

Major Area	Frequency	Percentage
Arts and Humanities	0.26	26
Social Sciences	0.31	31
Natural Sciences	0.13	13
Business Administration	0.11	11
Education	0.14	14
Engineering	0.05	5

It is sometimes useful to give the proportion or the percentage for each category, in addition to or instead of, the number of observations. By proportion, or relative frequency, for a given category we simply mean the fraction of all observations which fall into that category. For example, in the example above, the total number of observations is 250, and the number falling into the category of Arts and Humanities is 65. Thus the proportion, or the relative frequency, for that category is 65/250 = 0.26. The percentage for a category is the proportion multiplied by 100. Table 21 shows the same data as above, but in percentages. The relative frequencies are very useful if one is attempting to compare two sets of data when the total number of observations in each set is not the same.

MULTIVARIATE NOMINAL TABLES

If the observations are multivariate in nature--that is, if more than one characteristic is being measured on each observation--it may require somewhat more effort to construct tables. Table 22 is a three-way table for a nominal variable. This three-dimensional table shows frequencies for each of three measured characteristics: sex, religious preference, and political affiliation. Such tables often are referred to as contingency tables.

TABLE 22 -- MULTIVARIATE NOMINAL TABLE

Classification of 1,500 Registered Voters by Sex, Religious Preference and Political Affiliation

	Hindu	Moslum	Other Religion	No Religion	
Male					
Conservative	112	125	16	38	291
Liberal	81	157	12	43	293
Independent	39	26	11	75	151
	232	308	39	156	735
Female					
Conservative	143	146	21	28	338
Liberal	82	161	9	32	284
Independent	39	51	17	36	143
	264	358	47	96	765

With regard to multivariate data, it is sometimes the case that, in addition to having a table that shows all categories for all characteristics measured, we may also wish to summarize certain of the characteristics in a table having fewer dimensions. For example, from the three-dimensional Table 22 we could construct three different two-dimensional tables, merely by "summing over" one of the characteristics at a time. Table 23 shows the marginal distribution resulting when we no longer classify according to sex. The number of individuals in each political-religious class was obtained by "summing" over the sex characteristic. Table 24 shows the data displayed as a univariate distribution for the political affiliation characteristic alone. Again the counts, or frequencies, for this marginal distribution were obtained by summing over the categories of the characteristics being eliminated. Any of these marginal tables may display proportions or percentages rather than actual counts if this procedure appears to be useful. The relative frequencies, as well as the frequencies, are given in Table 24.

TABLE 23 -- BIVARIATE TABLE BY POLITICS AND RELIGION

Classification of 1,500 Registered Voters by Religious Preference and Political Affiliation

	Hindu	Moslum	Other Religion	No Religion	
Conservative	255	271	37	66	629
Liberal	163	318	21	75	577
Independent	78	77	28	111	294
	496	666	86	252	1500

TABLE 24 -- UNIVARIATE TABLE BY POLITICS ONLY

Classification of 1,500 Registered Voters by Political Affiliation

Political Affiliation	Frequency	Relative Frequency
Conservative	629	0.42
Liberal	577	0.38
Independent	294	0.20

In addition to determining the marginal distributions, obtained by merely adding through "columns" or "rows" of a table, it is sometimes useful to construct tables of conditional distributions for multivariate variables. A conditional distribution for a variable characteristic is obtained by counting and categorizing only those observations that meet a specific requirement or condition.

Conditional distributions may be either univariate or multivariate. For example, from Table 22 we may form the bivariate conditional distribution for political affiliation and sex, under the condition that we consider only those persons who are classified as Moslum. This distribution is given in Table 25. As with other tables, the conditional distributions may be presented either in the form of relative frequencies or in the form of percentages. Table 5(b) is a univariate conditional distribution of political affiliation, used under the condition that we are considering male voters only. The entries in Table 26 are given both as frequencies and as relative frequencies. (Note that the proportions of the conditional distribution for political affiliation are not the same as the proportions of the marginal distribution for that same characteristic given in Table 24.) Thus it can be seen that multivariate tables can be "built up" from appropriate sets of conditional tables.

TABLE 25 -- BIVARIATE TABLE BY RELIGION AND SEX

Political Affiliation and Sex for 666 Moslum Land Owners

Sex	Conservative	Liberal	Independent
Male	125	157	26
Female	146	161	51

TABLE 26 -- UNIVARIATE TABLE FOR MALE VOTERS

Political Affiliation for 308 Registered Male Voters

Political Affiliation	Frequency	Relative Frequency
Conservative	125	0.41
Liberal	157	0.51
Independent	26	0.08

ORDINAL SCALE TABLES

Tables for ordinal variables often are simply listings of how the observations were ranked or put in rank order. By rank order we mean an arrangement of the observations in an "increasing" (or a "decreasing") order. If the observations include assignments of ranks by more than one observer, then either (a) each of the various rank orders would be given separately, by listing individual observer rankings, or else (b) each separate observation would be listed together with the ranks it had received from the various observers. If a number of observations are put into a rank order consisting of only a few levels, the resulting table will look much like one for nominal data, except that in the ordinal scale table the classifications for the variable will be ranked.

Variables may possess both ordinal and nominal characteristics. For example, Table 27 shows a bivariate distribution with measurements on plans for university attendance (ordinal) and on sex (nominal). The marginal distribution for plans for university attendance (the "total" column) is a univariate table for ordinal scale data alone.

TABLE 27 -- TABLE OF ORDINAL SCALED DATA

Attitude Toward University Study of 700 High School Students

Plans for University Attendance	Male	Female	Total
Will not attend	52	93	145
Probably will not attend	120	82	202
Probably will attend	138	74	212
Will attend	90	51	141

INTERVAL SCALE TABLES

Considerable work often is required to achieve a good tabular presentation of interval data. The usual situation is that the numerical observations do not fall naturally into distinct classes. As an illustration, consider the 100 observations given in Table 28. The numbers presented there represent a sample of 100 different ways, with the characteristic reported being the number of ships. As they are recorded in the table, the data certainly cannot be analyzed conveniently. It is difficult for one to gain any meaningful impression with regard to the most or least common numbers, the numbers at which concentrations of observations might be present, and so on. Granted, ranking the 100 observations would be helpful, but a more concise and efficient method for ordering the observations would be to construct a frequency distribution. The procedure to be followed in doing so consists of grouping the data into intervals of numerical values rather than listing each value uniquely. Table 29 is an example of a frequency distribution with interval data. The tables given previously for nominal data also contained frequencies; but, in those cases, the category designations do not have a numerical order or numerical grouping. (Although tables for nominal and ordinal data may also be formed by grouping categories and are sometimes called frequency distributions, the term is usually reserved for interval data tables.)

TABLE 28 -- TABLE OF UNSORTED RAW DATA

Number of Merchant Ships Anchored at Sevastopol' on 100 Different Summer Days (Data Collected 1977-1980)

33	46	62	49	40	53	59	42	48	34
63	30	49	36	51	19	27	38	24	41
25	48	20	50	39	41	33	31	41	27
40	42	39	31	47	46	54	56	35	48
16	58	48	40	57	25	43	40	37	43
49	35	46	61	33	45	55	52	43	39
41	44	23	37	41	37	42	45	50	54
35	38	32	41	53	41	57	32	48	45
62	40	55	45	37	57	49	56	54	29
26	54	49	36	50	39	43	38	44	32

TABLE 29 -- TABLE OF FREQUENCY DISTRIBUTION

Number of Ships in Sevastopol' on 100 Different Summer Days

Number	Tally	Frequency
15-19	//	2
20-24	///	3
25-29	///// /	6
30-34	///// /////	10
35-39	///// ///// ///// /	16
40-44	///// ///// ///// ///// /	21
45-49	///// ///// ///// ///	18
50-54	///// ///// /	11
55-59	///// ////	9
60-64	////	4
	Total	100

In order to construct a frequency distribution for the observations of Table 28, we would first have to decide how many classes (groups) we should have and what the end points of those classes should be. Usually, frequency distributions are constructed containing from six to fifteen classes. Employing fewer than six classes generally results in one's losing too much information, because of the large intervals that often occur. At the other extreme, having more than 15 classes rarely is necessary to give a reasonably complete picture of what the observations really are like. We find that, for our data the smallest observation is 16 and the largest is 63. The range of 63 - 16 = 47 can be covered conveniently by 10 classes, each embracing a five-year period. (We could have used more or fewer classes, to be sure, but having the lower end points for the classes stated in multiples of five has some appeal.) A convenient method of determining the frequency for each class--that is, of determining how many observations fall into each of the designated classes--is to proceed through the list of recorded figures, making tallies.

The numbers we have used in our table to designate the classes are called the class limits. In Table 29, the first class listed has a lower class limit of 15 and an upper class limit of 19, the second class has a lower limit of 20 and an upper limit of 24, and so on. If there is a "gap" between the upper limit of one class and the lower limit of the next larger class, it is common to refer to the dividing point (usually taken equidistant between the two) as the class boundary. or real class limit. Referring to Table 29 once again, we find that the lower class boundary for the first class would be 14.5 (for we assume that if there had been a class lower than the first one it would have been formed similar to the others and thus have had a lower limit of 10 and an upper limit of 14) and that the upper class boundary for it would be 19.5. Continuing in the same fashion, the lower class boundary for the second class would be 19.5, and the upper class boundary for it would be 24.5, and so on. In some tables the values for class limits and class boundaries are identical. For example, this would have been the case in Table 29 if we had designated our classes as being 15 and under 20, 20 and under 25, 25 and under 30, and so on.

The class width (sometimes called simply the class interval) is the difference between the upper and lower boundaries of the class. Thus in table 8 the width of all the classes is five (not four). Additional terms used in conjunction with frequency distributions are class midpoint and, less commonly, class mark. The class midpoint is, as the word indicates, the value lying midway between the class limits, or, equivalently, midway between the boundaries. In Table 29 the class midpoints are 17, 22, 27, 32, and so on, which numbers can be computed in either of the following ways: (15 + 19)/1 = 17, (20 + 24)/2 = 22; or (14.5 + 19.5)/2 = 17, (19.5 + 24.5)/2 = 22, and so on.

While there are certain advantages to having all the classes be of the same width, this is sometimes not a feasible practice. Equal class widths become undesirable when a large proportion of the observations are concentrated within a relatively short interval. Family incomes furnish us with an example of this type of distribution. Even if we were to use as many as twenty classes, we would find that an extremely large proportion of our total number of families would fall into the lowest class, in that well over 95 percent of the cases under investigation are under $50,000 per year, whereas the upper limit for the rest is much higher. Table 30 illustrates a frequency distribution having unequal class widths. Moreover, this distribution contains an open end class (that is, a class for which only one of the two limits is specified).

TABLE 30 -- FREQUENCY DISTRIBUTION WITH UNEQUAL CLASS WIDTHS

1993 Populations of European Urban (Incorporated) Areas

Population	Number of Urban Areas
0 and under 2,500	2,727
2,500 and under 5,000	8,038
5,000 and under 10,000	12,924
10,000 and under 25,000	21,415
25,000 and under 50,000	17,848
50,000 and under 100,000	12,724
100,000 and under 250,000	4,286
250,000 and under 500,000	742
500,000 and under 1,000,000	267
1,000,000 or more	29

Adapted from: The Book of Great Misinformation

As is true in the case of nominal scale data, we may also list proportions or percentages, rather than frequencies, when making up a table for interval data. The choice will be dependent upon the uses that might be made of the data. In addition, we may have multivariate tables for interval scale data alone or multivariate tables combining interval scale characteristics with ordinal or nominal characteristics. Table 31 is a example of a bivariate table with both of the characteristics given on an interval scale.

TABLE 31 -- BIVARIATE INTERVAL DATA TABLE

Years of School Completed by Age for Soviets
25 Years Old or Older (In thousands)

Age in Years	Years of Schooling						
	0-4	5-7	8	9-11	12	13-15	16 or more
25-29	153	390	570	2,059	6,076	2,324	2,351
30-34	163	390	616	1,931	5,164	1,500	1,814
35-44	593	1,345	1,688	4,103	9,931	2,644	3,032
45-54	846	1,833	2,726	4,301	8,930	2,444	2,420
55 & older	1,801	5,624	8,805	6,206	8,456	2,870	2,987

Adapted from: Imagination, the Art of Wild Harry.

CUMULATIVE FREQUENCY DISTRIBUTIONS

Another tabular form that is useful in presenting ordinal or interval data is the cumulative frequency distribution, or simply, the cumulative distribution. This type of table lists the number of observations that fall above or below certain given values. Table 32 represents a "less than" cumulative distribution for the data of Table 29. The cumulative frequency column indicates the number of observations that fall below each of the "less than" values. (For example, from Table 29 we see that the lowest four classes contain 2 + 3 + 6 + 10 = 21 observations. Thus in Table 32 the cumulative frequency entry for "less than 35" is 21.) As is the custom, we have made use of lower class limits in Table 32 however, we could also have used lower class boundaries. Relative cumulative distributions and percentage cumulative distributions can be constructed by calculating the proportion or the percentage of the total number of observations falling either above or below the stated limits. That is, relative cumulative frequencies can be obtained by dividing each of the cumulative frequencies by the total number of observations.

TABLE 32 -- CUMULATIVE FREQUENCY DISTRIBUTION

Number of Ships in Sevastopol' on 100 Different Days

Number	Cumulative Frequency
Less than 15	0
Less than 20	2
Less than 25	5
Less than 30	11
Less than 35	21
Less than 40	37
Less than 45	58
Less than 50	76
Less than 55	87
Less than 60	96
Less than 65	100

Cumulative frequency distributions of an "or more" form can also be constructed. In this case, using the data of Table 29 we would label the rows as being 15 or more, 20 or more, 25 or more ... 60 or more, and 65 or more, and would list the cumulative frequencies as being 100, 98, 95, . . ., 4 and 0. "Or more" forms may also make use of relative frequencies or percentages. As a descriptive device, the cumulative frequency distributions show at a glance how many observations occur below (or above) any particular class limit.

GUIDELINES FOR MAKING TABLES

1. Tables should be fully self-explanatory. All information needed to comprehend a table's content should be contained in the table's title, column and row headings, and (if needed) footnote(s).

2. Table values should be grouped to provide maximum clarity and ease of interpretation. Spacing between groups usually is better than separating by lines.

3. The units of measurement for every table entry should be identified clearly.

4. Keep tables simple. Several simple tables are better than one complicated table.

PROBLEM SOLVING TECHNIQUES

5. Entries of dashes (--), etc., should be used only to indicate missing data. Entries having a measured value of zero should be entered explicitly as zero.

6. Decimal entries should begin with a leading zero (0.---) not just a decimal point.

7. All entries having a common measure should have the same number of decimal places.

INDEX

INDEX

PROBLEM SOLVING TECHNIQUES

D

E

F

G

INDEX

INDEX

PROBLEM SOLVING TECHNIQUES

Q

INDEX

PROBLEM SOLVING TECHNIQUES