NOT FOR QUOTATION
WITHOUT PERMISSION
OF THE AUTHOR

THE METHOD OF APPLIED SYSTEMS ANALYSIS: FINDING A SOLUTION

W. Findeisen
E. S. Quade

April 1980
WP-80-62

Working Papers are interim reports on work of the International Institute for Applied Systems Analysis and have received only limited review. Views or opinions expressed herein do not necessarily represent those of the Institute or of its National Member Organizations.

INTERNATIONAL INSTITUTE FOR APPLIED SYSTEMS ANALYSIS
A-2361 Laxenburg, Austria

April 28, 1980

FOREWORD

The International Institute for Applied Systems Analysis is preparing a *Handbook of Systems Analysis*, which will appear in three volumes:

- *Volume 1: Overview* is aimed at a widely varied audience of producers and users of systems analysis
- *Volume 2: Methods* is aimed at systems analysts who need basic knowledge of methods in which they are not expert; the volume contains introductory overviews of such methods
- *Volume 3: Cases* contains descriptions of actual systems analyses that illustrate the methods and diversity of systems analysis

Volume 1 will have ten chapters:

1. The context, nature, and use of systems analysis
2. Applied systems analysis: a genetic approach
3. Examples of systems analysis
4. The method of applied systems analysis: finding a solution
5. Formulating problems for systems analysis

6. Generating alternatives for systems analysis

7. Estimating and predicting consequences

8. Guidance for decision

9. Implementation

10. Principles of good practice

To these ten chapters will be added a glossary of systems analysis terms and a bibliography of basic books in the field.

Drafts of this material are being widely circulated for comments and suggestions for improvement. In addition to responding to such interventions, the task of detailed coordination of the chapters—prepared separately by several authors—has yet to be carried out. Correspondence about this material should be addressed to the undersigned.

This Working Paper is the current draft of Chapter 4.

A word about the format of this Working Paper. In order to make the text of each chapter easily amended, it has been entered into the IIASA computer, from which the current version can be reproduced in a few minute's time whenever needed. This Working Paper was produced from the version current on the date shown on each page.

Hugh J. Miser
Survey Project

April 28, 1980

CONTENTS

THE METHOD OF APPLIED SYSTEMS ANALYSIS: FINDING A SOLUTION

W. Findeisen, and E.S. Quade

4.0 INTRODUCTION

Usually, for a systems analysis to be undertaken, someone must have a problem, that is, he must be dissatisfied with the current or anticipated state of affairs and want help in bringing about a change for the better. Systems analysis can almost always provide some of this help, even if it does no more than present relevant information. However, the goal most frequently sought for systems analysis is to discover a course of action that will bring about a desired change for the better—that is, a course that can be judged to be most advantageous by those who have the authority to act.

Although discovering ameliorative solutions is its first task, systems analysis can frequently be used to help bring the solutions it discovers to acceptance by both the responsible policymakers and the people affected. In addition, after a solution is accepted, systems analysis can be applied during the process of implementation to help prevent the chosen course of action from being vitiated by adverse interests, misinterpretations, or unanticipated problems.

This chapter presents the methods of systems analysis in so far as they relate to discovering better solutions; advice as to how it can be used to win acceptance or to aid in implementation is postponed to later chapters. The problems of winning acceptance for a course of action and then implementing it must, nevertheless, be considered during the process of seeking and evaluating solutions, for a proposed course of action that is not acceptable to those who must adopt it, or that cannot be implemented for political or other reasons, cannot be a solution.

Although a systems analysis may be carried out without a specific decisionmaker in mind, this is not the usual case. This chapter discusses the methodology as if the analysis were being carried out for a single decisionmaker who commissioned it. This decisionmaker is assumed to be an individual who wants to make his decisions by taking into consideration the probable consequences of each of his possible courses of action; as a simple extension, we can also consider the single decisionmaker to be replaced by a relatively small group. The analyst's basic procedure is to determine what the decisionmaker wants, search out the alternatives that are available, work out the consequences that would follow the decision to adopt each of the alternatives, and then either rank them in terms of their consequences according to a criterion specified by the decisionmaker or present them to the decisionmaker for ranking in some framework suitable for comparison.

In reality, the decisionmaking situation rarely fits this paradigm; the persons for whom a study is done are usually no more than key participants in a decisionmaking process, who use the results of the analysis as evidence and argument to bring others to their point of view. Although this more realistic decisionmaking situation may introduce complexities in executing some of the steps, it is our view that no major change in the basic analytic procedure is required. Throughout this chapter we stick to the basic—or unsophisticated—view of the decisionmaking situation. For further views on the decisionmaking paradigm see Allison [1], Rein and White [2], Lynn [3], Nelson [4].

As an example, assume that a legislative committee wants to propose legislation to increase highway safety. They are willing to consider measures of three types: a requirement for devices to make the use of seat belts mandatory, lowering the maximum speed limit and enforcing it more strictly, and establishing higher standards for issuing driver's licenses. They ask the legislative analyst to carry out a systems analysis.

In the simplest systems-analysis approach (which is essentially identical to the traditional decisionmaking paradigm) it is useful to consider the problem in terms of the following elements:

Objectives. What the decisionmaker desires to achieve. In the example, it is increased highway safety, a concept that the analysis must make more precise.

Alternatives. The means by which it may be possible to achieve the objectives. In the example, there are three kinds of alternatives.

Costs. The cost of an alternative is the totality of things or actions that must be given up to acquire the alternative, including money, the use of personnel or facilities for other purposes, and so on. For example, stricter enforcement of the speed limit would require more police officers, who must be hired and trained or taken from other tasks; in either case the action would result in a cost to be associated with any speed-control alternative.

Performance Scales. A performance or effectiveness scale is a device for indicating the extent to which an objective is attained. It provides a tool for evaluating the performance of alternatives in achieving the objective. For example, it can be agreed to measure the increase in highway safety by the decrease in annual traffic fatalities.

Performance. The performance or effectiveness of an alternative is the position on the scale it achieves.

Criterion. A rule for decision that specifies in terms of performance and cost how the alternatives are to be ranked. A common one is to rank the alternatives in decreasing order of performance for fixed cost.

Models. The models are means by which the performance and costs, and in most cases the other consequences or impacts associated with adopting and implementing the different alternatives, are estimated or predicted. Further discussion may be found in Quade [5] and [6].

Different models are required for different alternatives and for different purposes; a model to estimate the monetary costs of doubling the strength of the highway patrol differs from a model for predicting the effect the presence of this increased force on the highways will have on traffic fatalities. A model is made up of the factors relevant to the problem and the essential cause-and-effect relations among them. It may consist of a set of tables, a series of mathematical equations or a computer program, a physical simulation (but rarely for systems analysis), or merely a mental image of the situation in the mind of the analyst made explicit with a sequence of logical arguments.

The objective and systematic approach by means of an explicit model is needed for predicting the impacts because, in most systems-analysis problems, the factors are so numerous and their interrelations so complex that intuition is not good enough. Some safety measures, for instance, have counterintuitive effects: certain crash barriers reduce fatalities but increase accidents. Others have interdependencies that strongly affect their joint performance: an energy-absorbing bumper would appear to save more lives if it were installed alone than in combination with a shoulder harness [7].

In our example, an early problem for the analyst is to find a way to turn the vague goal of "increased highway safety" into something of a more operational character, or at least settle on a way to measure it. One possibility might be to use the reduction in the annual number of fatalities as a measure of performance; another might be to use the reduction in the annual (monetary) cost of highway accidents. Unfortunately, this choice may affect critically how the alternatives are ranked. For instance, while strict enforcement of the speed limit may reduce fatalities, a serious consequence of high-speed collisions, it may have little effect on the number and cost of "fender-bending" accidents, which are numerous, while more stringent requirements for a driver's license may reduce them significantly.

Another early task for the analyst is to examine the alternatives for feasibility. It may turn out, for example, that, in the current state of the art of automotive engineering, the alternative of automated seat belts is not feasible, say, owing to reliability considerations. Similarly, the analyst may be able to find out that the passage of legislation to lower the current maximum speed limit is not politically feasible. This alternative may then have to be reduced merely to stricter enforcement of traffic regulations, dropping any thought of lowering the maximum speed limit.

The analyst will also want to examine alternatives not on the original list--such things as better emergency ambulance service, eliminating grade crossings, changed car design, and others—that may promise lower fatalities at no greater cost.

In predicting the impacts associated with the alternatives, the analyst may have to use radically different means or methods. A model to show the effect of improved driving skills on fatalities can be considerably different from a model to predict the way a lower speed limit affects fatalities. On the other hand, predictions for both cases may be obtained statistically from experiences in other jurisdictions with similar driving conditions. Also, to compare alternatives, various different futures may have to be considered, with assumptions made about the effects of a petroleum shortage on automobile traffic, changing car design, population movement, and other exogenous factors beyond the decisionmaker's control that can affect the outcome.

One run-through of a problem is seldom enough; several cycles or iterations usually improve confidence in the results. For instance, it may be discovered that the impacts of certain alternatives that restrict automobile drivers produce effects that "spill over" onto entirely different groups of people, say those that ride public transportation, in ways that differ from alternative to alternative and were not anticipated when the alternatives were first formulated. Additional emergency medical services for traffic-accident victims, for instance, may increase the burden on the supply of doctors and hospital beds, and hence the analyst may have to enlarge the analysis to include aspects of the medical system.

With this background, we now turn to a more detailed and thorough description of the procedures we have suggested.

4.1 A FRAMEWORK FOR SYSTEMS ANALYSIS

Objectives, alternatives, choice

Analysis to assist someone (called here the decisionmaker) to discover his "best" course of action may, in general, be considered as an inquiry into these three basic questions:

1) What are the decisionmaker's objectives?

2) What are his alternatives for attaining these objectives?

3) How should these alternatives be ranked?

As defined earlier, the *objectives* are what a decisionmaker seeks to accomplish or to attain as a result of his decision, and the *alternatives* are the means available to him for attaining the objectives. Depending on the particular problem, the alternatives may be policies, strategies, systems designs, or actions. *Ranking* implies designating the alternative that is "best," considering the consequences of implementing the various alternatives, the objectives, and the values the decisionmaker puts on the outcomes.

The three basic questions expand into further questions when we consider that:

- to be able to identify the feasible alternatives, one must know, not only the objectives, but also the boundaries within which the decisionmaker is free to act, that is, the *constraints;*

- to compare and rank alternatives, one must predict the *consequences (impacts)* that are likely to follow from the choice of each alternative;

- to determine the consequences of an alternative, we need a *predictive* (cause-and-effect) *model* showing what will happen if the decisionmaker chooses the alternative, given a particular contingency, or alternative future *state of the world* on which the predicted consequences will certainly depend;

- to help the decisionmaker rank the alternatives, it may be necessary to determine his *value system* and possibly that of other parties affected by some of the consequences.

Because alternatives may differ radically, we may need a different model for each alternative. We may also use different models, from the very rough to the very precise, as we proceed in the analysis from the first assessments to final results.

A framework for analysis

Systems analysis to aid decisionmaking is a craft activity. The way in which a study is organized and performed depends on many choices by the analyst—called secondary decisions after White [8]—that are often based on little more than intuition. An approach that may produce valuable insights when used by one analyst may yield faulty or misleading conclusions when used by another. Nevertheless, every systems analysis will be composed of certain more or less typical activities that have to be appropriately linked to each other. From this point of view, we can present a first approximation to the systems-analysis process schematically as in Figure 1, where the main components are represented (other breakdowns are, of course, possible):

1) Formulating the problem.

2) Identifying, designing, and selecting the alternatives to be considered.

3) Forecasting the operational context or state of the world.

4) Model building and predicting the consequences.

5) Comparing and ranking outcomes.

These components encompass several additional activities, two of which are indicated in Figure 1: ***determining the constraints*** and ***determining the decisionmaker's values and criteria***. Among those omitted from our list, but needed for almost every analysis, are ***data collection and analysis***, and ***communication between analyst and decisionmaker***.

The solid lines in Figure 1 show the principal flows of information from activity to activity.

Iteration

In most investigations few of the component activities depicted in Figure 1 can be performed adequately in a single trial. Iteration is needed; that is, preliminary results, or even an incomplete version of the final result, may force the analyst to alter initial assumptions, revise earlier work, or collect more data. A decisionmaker, for instance, may not settle on his objectives until he has a good idea of what he can do, or he may want to impose additional constraints after he discovers what some of the impacts are.

Figure 2 shows some of the typical iterations in a systems analysis study.

One feedback loop is from the impacts (the consequences) to designing alternatives. By this loop one modifies or refines some alternatives, typically by adjusting their parameters, and eliminates others. The process of refinement through iteration

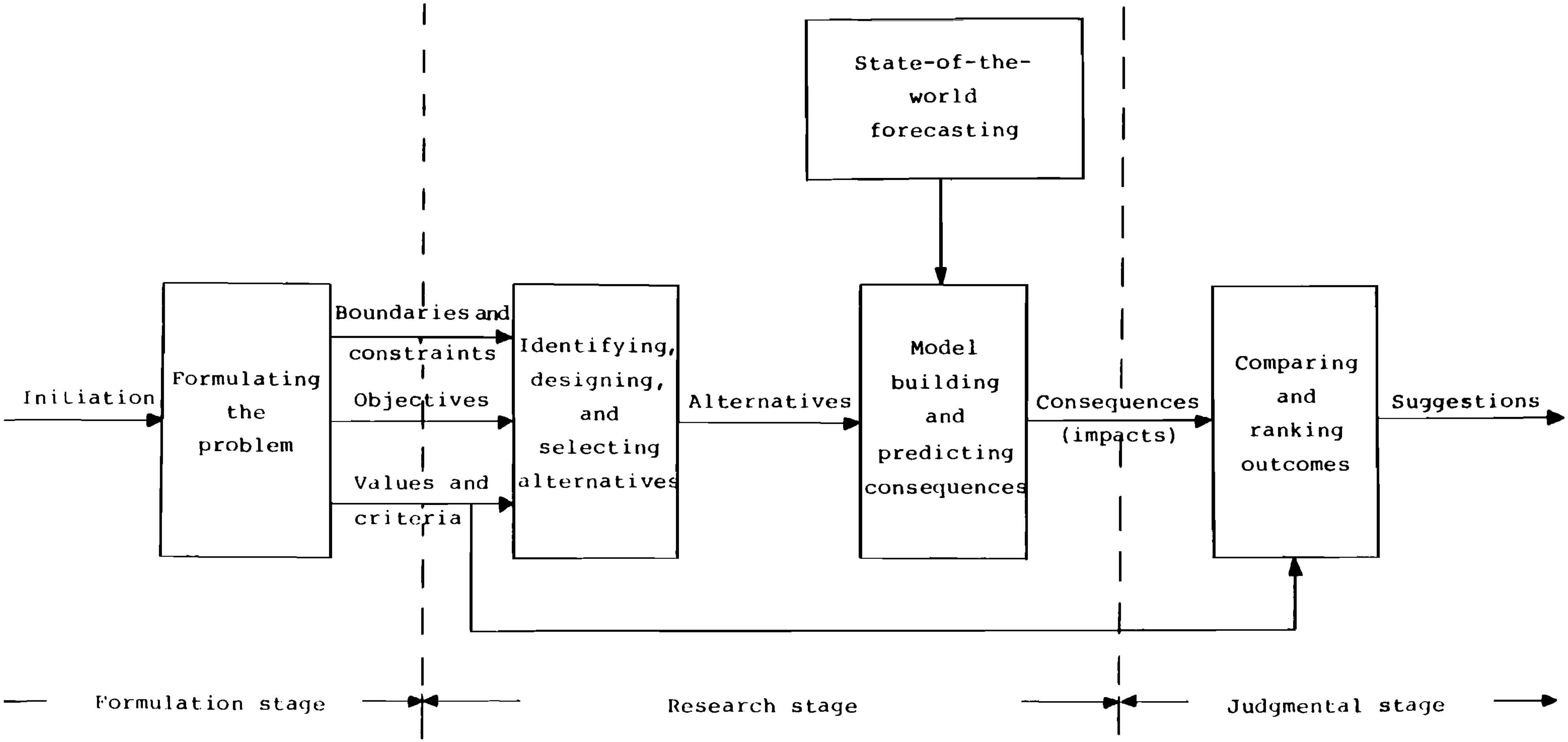

Fig. 1. The systems-analysis procedure

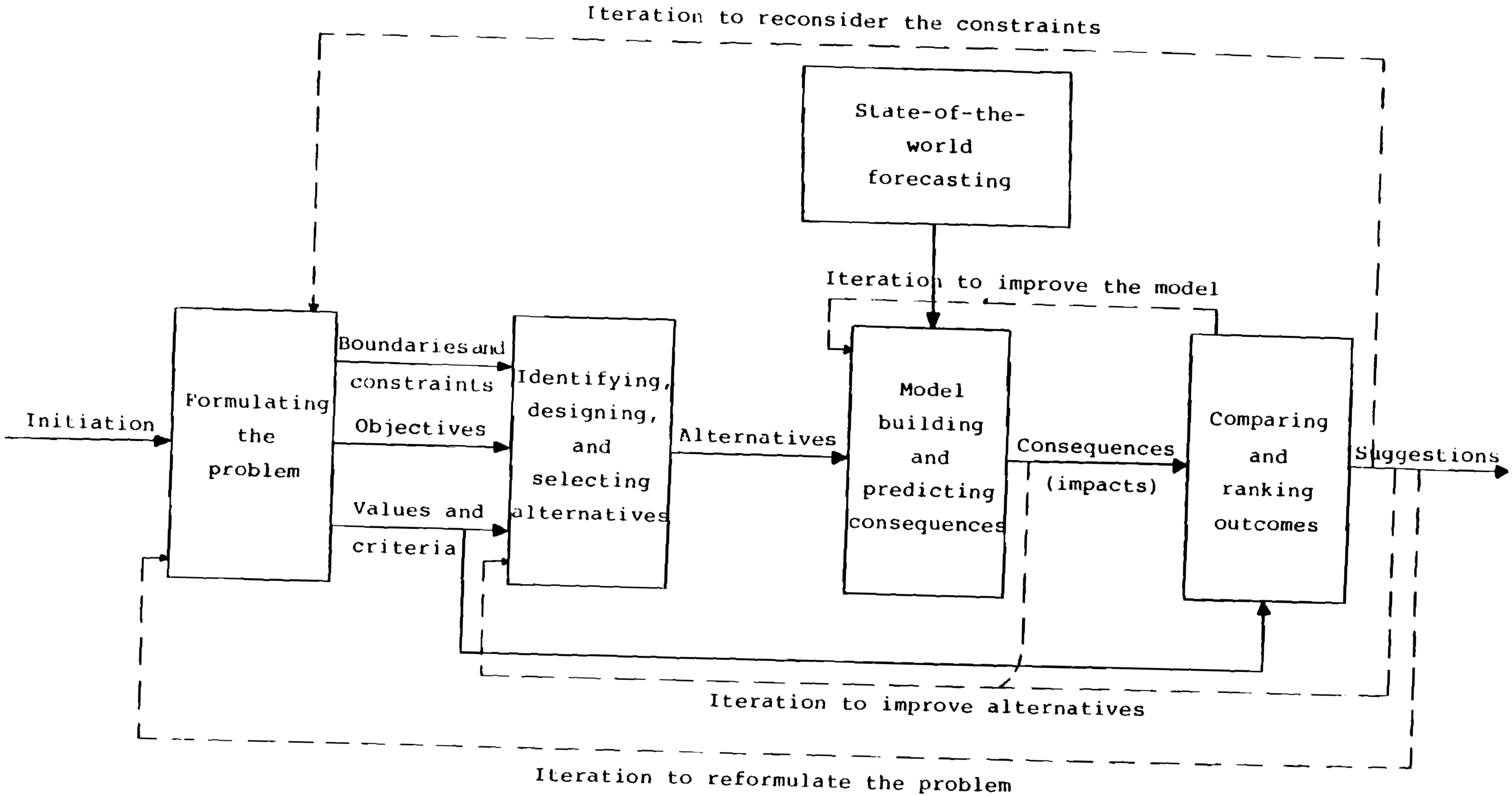

Fig. 2. The systems-analysis procedure with iteration loops

may be done separately for each alternative; it is sometimes based on a formal optimization procedure.

Another typical loop is the one from the model results back to problem formulation. This iteration is necessary because it is usually impossible to set the objectives and determine the constraints with any sort of precision before knowing something about the implications of the assumptions. Iteration may require redefining the alternatives; that is, we may have to design an entirely new set of alternatives.

Furthermore, we may be dissatisfied with the results obtained under our current assumptions and constraints. Iterations may be carried out to see what the "cost" of the constraint is, that is, how much more of the objective could be obtained if a constraint were changed. We may eventually negotiate removing, or softening, some of the constraints. If this is not possible, lowering the objectives of the decisionmaker may have to take place.

Another important purpose of iteration is to improve the model, a process that may actually simplify it, and thus lead to less detailed but more secure (robust) decisions.

Stages of analysis

There are many more linkages between the component activities of systems analysis than those shown in Figure 2. Despite this complex interdependence, it is convenient to discuss the procedure in three *stages*:

A. Formulation Stage

B. Research Stage, comprising

- Generating and investigating alternatives

- Determinating the consequences

- State-of-the-world forecasting

C. Evaluation (comparison and ranking) Stage

Usually, the analyst is not finished, even when iteration no longer brings significant improvement and the various courses of action open to the decisionmaker have been compared and ranked. As mentioned in the introduction, an analyst, although not necessarily the original one, may also be called on to provide assistance with additional tasks--securing adoption and implementation of the results, and, later, after the work of implementation has succeeded (or failed), evaluating the entire process.

Whether a course of action is feasible from the point of view of implementation is not the main question when the analyst is helping with the actual implementation; infeasible alternatives should, ideally, have been eliminated during the earlier stages of the analysis. It is rather that the final decision may not be adequate to instruct and motivate those who have to execute the decision and who may have their own ideas as to how to interpret it. There may also be considerations that are important for implementation, but which were not important to the choice between alternatives and which, in order to keep the problem workable, have not been spelled out in detail.

A decision may take so long to implement that changes in the state of the world different from any of those forecast in the analysis may require its modification. What was "the future" during the analysis becomes the present, and the analyst may be needed once again to modify a course of action that may now be partially obsolete.

In addition, the analyst may be called on to assist the decisionmaker to evaluate the progress of the implementation, for, by virtue of his previous studies of the problem and his knowledge of the cause-effect relations, he may be able to detect the reasons for discrepancies and deviations from the effects originally intended.

Communication

In this context it is worth while mentioning a very important factor in systems analysis: communication. Communication with the decisionmaker is vital, although communication with others is also needed. The decisionmaker's advice and judgments are indispensable at all stages of the analysis. Its results are much more likely to be accepted and used if the decisionmaker participates in their production. Throughout the procedure there should be a continuous dialogue between the analyst and the decisionmaker—including his staff. This dialogue influences the decisionmaker's attitude toward the problem even before the study is finished, and helps to make sure that the important facets of the real situation are considered.

One reason for constant communication is that the initial problem formulation can never be complete and all-inclusive. As mentioned above, partial results of analysis will modify the initial views, new questions will arise, and the preferences, constraints, and time horizons may change.

A constant flow of information to and from the decisionmaker and his staff will give them a sense of participating in the study and will mean that the results will not come to them cold, with a sense of shock—which can lead to their rejection.

Partial analysis

We shall characterize the *stages* of systems analysis, as well as the more important *component activities*, in more detail shortly.

Before doing this, however, we note that not every systems-analysis study contains every stage or component. Some studies may be useful even though they lack some of the steps in the very general schematic presentation in Figure 1; we refer to such studies as *partial analyses*. Here are some typical examples:

- Forecasts of the future state of the world, where no immediate action by a decisionmaker is contemplated; for example, econometric forecasts, which analysts are asked to provide for governments or large industrial companies.

- Impact analysis, i.e., determining all impacts, or even merely certain impacts, of a proposed course of action. For example, studies to determine the consequences of a particular technological development on the environment may involve no comparison or ranking.

- Decision analysis, that is, assistance in making a choice among a limited number of well specified alternatives, whose consequences are assumed to be known. Here the analysis merely provides a framework for ranking these alternatives. A typical instance is the choice of an industrial project from among several available alternatives, or a decision to buy equipment from competitive suppliers.

In these examples, not all of the component activities of a complete systems analysis are carried out by the analysts. On the other hand, there are cases, where all the activities are present, but where some of them need to be emphasized more than others.

Whenever a partial analysis is commissioned, the assumption is that the decisionmaker himself is providing the missing aspects through judgment or assumption. Thus, the need for good communication is particularly important here.

4.2 PROBLEM FORMULATION

Goals and difficulties

Generally speaking, problem formulation implies isolating the questions or issues involved, fixing the context within which these issues are to be resolved, clarifying the objectives, identifying the people to be affected by the decision, discovering the major operative factors, and deciding on the initial approach to be taken in the analysis. It is expected that problem formulation will provide, among other things:

(a) a preliminary statement of the objectives, and ways to measure their achievement,

(b) a specification of some promising courses of action, i.e., the alternatives,

(c) a definition of the constraints,

(d) an anticipation of the consequences, the measures of their value, and a definition of the criteria for choice.

Problem formulation should result in specifying its limits, what question is to be addressed, and what aspects of the real world are to be included, in what time frame, with what analytic resources.

During this formulation, the analyst must consider the analytic approach to be taken, which, of course, depends on the information and the type of problem. For example, the decisionmaker may have been assigned a fixed budget and desire to find the most effective available alternative, or, he may have a desired level of effectiveness he wants to achieve, in which case the objective of the analysis becomes identifying the least-cost mix of alternatives. Another possibliity is that progress is required in the correcting some undesirable condition, and the analytic objective is to discover the point at which the marginal benefits of corrective action become equal to the marginal costs, or to ascertain whether some proposed course of action yields a sufficiently high rate of return on the required investment to make it attractive.

As the study progresses and more information becomes available, the analytic approach may have to be modified.

In a sense, formulation is the most important stage of analysis, for the effort spent restating the problem in different ways, or redefining it, clarifies whether or not it is spurious or trivial, and may, indeed, point the way toward a solution.

Among the difficulties of problem formulation these usually stand out:

(i) No issue is isolated; every system is linked to other systems and it is thus part of a larger one. There is therefore a mutual dependence of the objectives, constraints, and consequences.

(ii) We cannot set the objectives firmly unless we know what can be achieved, that is, until we know—with reasonable accuracy—the results of analysis.

(iii) The objectives, as well as the measures of value and the criteria for choice, are highly subjective and depend on the decisionmaker's preferences, which may be both difficult to assess and varying over time. This applies, in particular, to high-level objectives, which are seldom stated in any sort of operational form.

For many reasons, the problem-formulation stage can be seen as almost a crude systems-analysis study in itself. It may involve a very broad range of inquiries into the hierarchies of objectives, the value systems, the various types of constraints, the alternatives available, the presumed consequences, how the people affected will react to the consequences, etc. A systematic approach to problem formulation through some fairly formal device such as an "issue paper" may be desirable; Chapter 5 describes this device and provides other information about problem formulation. One reason is that, until the problem has been defined and the issues clarified, it may not be clear that the study effort will be worth while.

Objectives

The objectives are what a decisionmaker seeks to accomplish or to attain by means of his decision, that is, by the course of action he decides to implement.

The analyst has to determine what the decisionmaker's objectives actually are; Chapters 5 and 8 give a more thorough discussion of the difficulties that are frequently encountered at this stage. For the present purpose, we state merely that an objective may be specified in a more or less general fashion, may be quantified or not quantified, and is usually a step in a *hierarchy of objectives;* one speaks about different *levels* of the objectives.

Often the levels of objectives differ according to the *time horizon.* For example, in economic planning, or in corporate planning, there is a hierarchy of short-term and long-term objectives that have to be consistent with one another.

The fear of setting up objectives that may prove to be inconsistent with higher-level, more comprehensive objectives, may lead a decisionmaker to specify an objective at too high a level to be helpful in the analysis. For one reason, the courses of action that are required to attain this higher level objective may not be his to choose.

It is the objectives that suggest the alternatives, for, to be considered an alternative, a course of action must offer, or appear to offer, some chance of attaining the objectives. As more information becomes available, the list of alternatives may increase or decrease.

Unless the objectives are correctly and clearly spelled out, the rest of the analysis will be misdirected—wrong and ineffectual alternatives will be proposed that do not favorably affect the problem that generated the analysis. To define objectives it is often helpful to call on several people not involved with the problem under analysis, particularly outsiders skeptical of what they think the decisionmaker is trying to do. Another possibility is to start by specifying a measure of performance that seems appealing and then examining the objectives it serves. In effect, one keeps trying to answer such questions as: What is the decisionmaker really trying to accomplish? What ultimate good result is desired? For example, what objective is really served by lowering the speed limit?

We would like to be able, for the sake of analysis, to *measure* the degree to which an objective will be attained by a course of action under consideration. For this reason, if the original objective cannot be quantified, one must often define a *proxy objective:* a substitute that points in the same direction as the original objective, but which can be measured. For example, "reduction of mean travel time" in urban transportation can be a proxy for "improved services."

If the degree to which the objective has been attained is measurable in some sense, one can set a *target value;* for example, "achieve an average travel time of 40 minutes". Often, to be more flexible, we prescribe an interval, for example, "achieve an average travel time of less than 45 minutes", which leaves more freedom for the choice of alternatives.

Another ambiguity that must be clarified is how the consequences of the course of action designed to attain a particular objective are related to it; for example, how do the various attributes of a transportation system as part of a program to improve the quality of urban life–time of travel, comfort, convenience, noise, air pollution, cost, some desirable, some undesirable—actually relate to this objective?

In many cases, the decisionmaker specifies *multiple objectives*. These objectives frequently contribute to a single higher-level objective, although we may not be able to measure how much.

An example of such a situation is "the quality of urban life," as a higher-level objective to which several component objectives, such as better housing, less air pollution, reduced travel times, less aesthetic discomfort, and others, contribute. If we cannot work out the relative contribution of each factor, we ordinarily seek alternatives that improve, in a measurable degree, all, or the majority, of the contributing component objectives, leaving the ultimate ranking to the decisionmakers.

Multiple objectives are usually *competitive*, i.e., an alternative designed to bring about maximum improvement in one of them is associated with a deterioration in some of others, because of limited resources or other constraints.

A reconciliation of the multiple objectives may present a serious problem, treated in Chapter 8 and in numerous publications (for example, Keeney and Raiffa [9] and Bell, Keeney, and Raiffa [10]).

Values and criteria

A course of action will have many consequences, some contributing to the objective, some detracting, with still others being side effects, that is, consequences that are neutral with respect to the objective, but possibly with productive or counterproductive implications. If we wish to say how good an alternative is we need a ***measure of value*** for each of its significant consequences. If we want, moreover, to be able to compare different alternatives in order to indicate a preference, we need *criteria* for ranking them in order of preference.

A measure of value is subjective. The same thing may be of different value to different people. In principle, it is the value or "desirability" for the decisionmaker that is important, because he will decide whether or not to take a given course of action. But, in all cases, the persons or groups the decisionmaker is serving, or who will be affected by his decision, must be considered.

For example, consider the air pollution to be caused by a future industrial plant. Assuming that no pollution standards or penalties exist, does this mean that the industrial manager can neglect pollution, although he knows the damage it will cause? Clearly, he cannot neglect pollution without a deliberate decision to do so, because the people affected may in one way or another, say through their influence on future standards, affect the profits of the plant. It is the duty of the analyst, in this case at least, to indicate the impact of pollution on those who will be affected, and somehow to transfer their subsequent dissatisfaction to the decisionmaker's balance sheet.

The values held by the decisionmaker, that is to say, the importance he attributes to the various impacts, determine the criteria for ranking the alternatives; hence the decisionmaker's values must be investigated at an early stage. We define a criterion as a "rule or standard by which to rank the alternatives in order of desirability." An example might be: "given a fixed task, rank the alternative first that can accomplish it at the least cost."

The values and criteria of interest are those of the decisionmaker. The aim of systems analysis, especially on public issues, is not to say what the decision *ought* to be; the analyst can only say that, given the criterion and his best knowledge about the decisionmaker's values, the alternatives should be ranked in a particular order. As soon as the analyst makes recommendations, based on his own values, as to what the decision should be, the analyst is abandoning his role as an analyst and becoming an advocate. This may be an appropriate role in some cases, but when assumed the analyst should make clear what he is doing.

More attention to the problem of criteria is given in section 4.6 on "Comparing and ranking alternatives" and in Chapter 8.

Constraints

Constraints are restrictions on the alternatives; they may be physical properties of systems, natural limitations, or imposed boundaries that do not permit certain actions to be taken. Thus, the constraints may imply that certain consequences cannot be obtained and that certain objectives cannot be achieved. The alternatives, consequences, and objectives that are not prohibited, directly or indirectly, by the constraints are referred to as *feasible*.

Some examples of possible constraints are: physical laws, natural-resource limitations, available manpower, existing legislation, accepted ethics, allocated investment money.

There are two main questions related to constraints in the analysis:

(i) What *are* the constraints?

(ii) What is their *influence*, i.e., what actions, and hence what consequences and objectives, are feasible?

Some answers to the first question will be discovered during problem formulation, but not all constraints may be revealed in the initial formulation; some may be discovered at later stages and others not until after implementation has started.

Finding answers to the second question is an essential part of the analyst's task. In fact, the question of feasibility is an important, if not dominant, component in systems analysis, and usually a difficult one to deal with. An investigation of the feasibility of actions or objectives is referred to as *feasibility analysis*.

There are many different kinds of constraints. Some are permanent and can never be violated (physical laws, global resources). Others are binding in the *short run,* but may be changed by the passage of time or by decision (e.g., legislation). Still others are arbitrary, set by the political situation or merely by the decisionmaker's tastes.

There are different constraints at different levels of decisionmaking. Usually the smaller the scale of a problem is, the more constraints are imposed on it in an arbitrary rather than in a natural and objective way. For example, an analysis of alternative urban transportation systems would have to consider a cost constraint, air and noise pollution standards, and perhaps also an employment constraint. All these are constraints imposed by decisions made at a higher level, usually of the resource-allocation type, and not directly by the available resources.

Depending on their character (objectively existing, or imposed by a decision) the various constraints are treated in essentially two different ways. Some constraints are *rigid* or *unquestionable;* to this category certainly belong the constraints of natural laws and natural resources. We have already indicated, however, that the latter are rigid only at a relatively high level. For a city, or an industrial plant, the resource constraints are often the result of an allocation decision and may therefore be considered *elastic* or *negotiable*. By elastic or negotiable constraints we mean ones that may, in principle, be changed by a higher-level decision if the analysis provides a good case for the change.

Providing the case may consist, for example, in showing how much more of the objective can be gained if the given constraint is changed by one unit. A calculation of this kind is called *marginal analysis*. It may happen, for example, that a slight lowering of the standard of admissible pollution would cause a substantial reduction in the cost of producing an industrial product.

Marginal analysis determines a price of the constraint; we should not forget, however, that it is, for example, the price or worth of the constraint to the polluting party, not to those who are being polluted.

What has been said so far about the constraints by no means applies only to constraints of a quantitative nature. Political and cultural constraints may have to be considered in certain analyses. Constraints of any kind may be divided into the categories of short-run and long-run constraints and into rigid and removable (or negotiable) constraints. There will also be different constraints at different levels of decision.

As already said, it cannot be expected that all constraints, and much less so the feasible sets that result from the constraints, will be revealed at the initial stage of problem formulation. Nevertheless, it is important to define at least the most influential constraints at this stage. With respect to those resulting from higher-level decisions, it is desirable to get some feel as to how firm these constraints are and, in particular, whether they are defined and definite for the whole time horizon. Otherwise, the analysis may investigate actions or alternatives that will be entirely inappropriate.

Constraints imposed by the client or decisionmaker are in the same category as goals, for it is doubtful if any real distinction can be made [11]. Goals, in fact, are little more than constraints selected for special attention because they motivate the decisionmaker or because it is convenient for the analysis.

A discussion of some of the difficulties in problem formulation may be found in Majone and Quade [12].

4.3 GENERATING AND SELECTING ALTERNATIVES

Characteristics

It can hardly be overstressed that generating alternatives is, in systems analysis, an exercise of creativity and imagination appropriately tempered by a thorough and broad knowledge of the issues. The alternatives that have to be considered in a particular case may be wide-ranging and need not be obvious substitutes for each other or perform the same specific functions. Thus, for example, education, recreation, family subsidy, police surveillance, and low-income housing (either alone or combined in various ways) may all have to be considered as possible alternatives for combating juvenile delinquency. In addition, the alternatives are not merely the options known to the decisionmaker and the analysts at the start; they include whatever additional options can be discovered or invented later.

The set of potential alternatives initially includes all courses of action that offer some chance of attaining or partially attaining the objectives. Later, as the constraints are discovered and applied, the set is reduced. The set of alternatives considered usually includes the "null" alternative, the case of no action, if only for the purpose of comparison.

In most cases, a number of alternatives are explicitly suggested by the decisionmaker, i.e., they are defined by a more or less detailed enumeration of their specific characteristics. For example, to improve urban transportation, both a subway to cover certain regions with stations and surface connections to outlying districts and a bus line with routes and schedules can be specified as alternatives to be considered.

Desired properties

The alternatives that survive to the final ranking will have to be feasible. They will also have to go a long way toward meeting the decisionmaker's objectives. But, apart from feasibility and meeting stated objectives, such as a desired carrying capacity for an urban transportation system and a wish to keep the mean travel time as low as possible, there are other vital features that, while they may not be explicitly stated as objectives, should be considered in creating and evaluating the alternatives.

One of these, an almost indispensable feature of an acceptable alternative, is its *insensitivity (robustness)*, measured by the degree to which attainment of the objectives will be sustained despite disturbances encountered in normal operation, such as varying loads, changing weather conditions, etc. In urban transportation, insensitivity could mean, for example, that the average travel time does not greatly increase even when the peak-hour load and street traffic are increased by 25%, 50%, or more.

Another feature important for many applications is *reliability*, which is the probability that the system is operational at any given time, as opposed to being out of order. In some cases, it is important for the proposed system never to fail; in others, that it not fail for a time longer than some threshold value; and in still other cases, a failure is tolerable if it can be repaired quickly; this feature, in turn, brings us to the question of maintenance and, consequently, logistics.

One says that a system is *vulnerable* if damage or failure of an element causes considerable trouble in meeting the objectives (vulnerability does not mean, or does not necessarily mean, complete failure). In the urban transportation example, a bus system is vulnerable to snow storms.

Flexibility is a property exhibited by an alternative designed to do a certain job when it can be used for a modified, or even an entirely different, purpose. It is important to have a flexible alternative when the objectives may change or when the uncertainties are very great. For example, for transferring fuel, rail transportation is more flexible than pipelines.

In addition, each alternative that survives the other hurdles must be examined with the problems of implementation in mind. Some alternatives will be easier to implement than others; those impossible to implement must be eliminated and the cost of implementation associated with each of the others must be taken into account.

Generating alternatives is above all a craft or art, an exercise of imagination, creativity, criticism, and experience. It is the diversity of alternatives, so often encountered in systems analysis, that calls for creativity and ingenuity rather than for a deep knowledge of formal tools. Therefore, what we say below can only be a loose guideline, a framework, which may be of assistance in some cases and useless in others.

Whenever a diversity of means exists to achieve the objectives, generating and selecting alternatives are best done in steps or stages. Initially, it is appropriate to consider a fairly large number of possiblities as alternatives; any scheme that has a chance of being *feasible* (that is, not likely to violate the constraints), and of *meeting the objectives* should be investigated. At the beginning, it is good to encourage invention and unconventionality; foolish ideas may not be foolish when looked at more closely. It may often be advisable to reach beyond the less rigid constraints, to broaden the scope of the study outside the limits that were initially set forth by the client.

The many alternatives that can be considered initially cannot all be investigated in detail. It would be too costly and, above all, excessively time-consuming. Some kind of screening, based on expert judgment, evidence from past cases, or simple models, can often be used to select a few of the alternatives as more promising for the next stages of investigation. It may, for example, be possible to reject some alternatives by *dominance:* i.e., because another alternative exists that is better in at least one aspect and equally good in all the remaining aspects considered.

The stages that follow the initial scrutiny should involve an increasing amount of quantitative assessment. At first, the assessment of the consequences of each alternative may still miss many details, but it should be adequate to permit rejecting a fair percentage of the original alternatives on the ground that the other cases are more promising.

The last stage of the selection procedure should investigate relatively few alternatives, but in considerable detail. These alternatives should be serious candidates for implementation. At this stage every effort should be made to assess each alternative as accurately as possible, and each one may have to be fine-tuned, to yield the best results possible.

As can be seen, we favor a precedure of step-by-step rejection of alternatives rather than one of focusing on selecting the best alternative in a single operation. This procedure has some rationale; first, the alternatives that are shown to be infeasible can be rejected (irrespective of what they promise in terms of benefits); next, the alternatives that can be shown to be markedly sensitive or vulnerable can be rejected, etc. It is, in many cases of judgment, easier to agree on rejection than to agree on positive selection.

Fine tuning alternatives

We use the term fine tuning to refer to determining the details of an alternative. Less detail may be needed for making a choice among the widely different types of alternatives, more detail may be necessary before selecting one of a selected type for implementation. At the implementation stage, systems analysis overlaps with "systems design" or "systems engineering," where—for example, for an industrial plant—the job is to determine all specifications for the consecutive design of the particular parts of the plant.

Fine tuning is an activity that may, in appropriate cases, make good use of mathematical models. The problems are usually well defined when fine tuning is appropriate and setting the details may be ideal for formal procedures for optimization, such as linear programming.

4.4 DETERMINING THE CONSEQUENCES

Future and uncertainty

An important analytic task is to predict the *consequences* (also referred to as impacts, effects, or outcomes) of each alternative that is being considered. The task is difficult because of uncertainty, particularly with respect to the future state of nature or the context in which the alternatives are to operate.

With the future in mind, assessing a course of action involves answering two questions:

(i) What will happen as a result of this action?

(ii) What will happen without this action?

Neither of these questions can ever be answered with certainty, because both involve one or more forecasts of future conditions, i.e., of the future states of the world or at least the segment of the world being considered in the study. These forecasts will often have the form of multiple scenarios.

It is essential to ask a question related to the probabilistic properties of these forecasts:

(iii) How certain are the answers to (i) and (ii) that the analysis can supply?

The last question may be split into various subquestions important for a given case, e.g., what is the range of likely outcomes of the action? Is there a possibility, even a very unlikely one, that the action's consequences will turn out to be very undesirable?

Listing the consequences

A particular alternative will have a large number of consequences. Some of these are *benefits*, things that one would like to have and which contribute positively to attaining the objectives; others are *costs*, negative values, things that one would like to avoid or minimize. Some of the consequences associated with an alternative may have so little apparent effect, positive or negative, on attaining the desired objective, that they are not considered in the analysis; they are referred to as *externalities*. Some of these, however, may affect or spill over on the interests of other groups of people or other decisionmakers, who in turn may affect the decision by making their objections known to the analysts or through pressure on the decisionmakers. It may therefore become necessary, in the course of analysis, to broaden the study to introduce the spillovers (which were previously externalities) into the revised analysis.

Broadening of the study can also change the judgment of what is a cost and what is a benefit. A new investment is a cost to the industrial company; but it may be a benefit from a regional or national point of view, if it helps reduce unemployment.

In the narrow sense, *costs* are the resources required to implement an alternative. In the broader sense, costs are the "opportunities foregone"—all the things we cannot have or do once we have chosen a particular alternative. Many, but by no means all, costs can be expressed adequately in money or other quantitative terms. Others cannot. For example, if the goal of a decision is to lower automobile traffic fatalities, the delay caused to motorists by schemes that force a lower speed in a relatively uncrowded and safe section of road will be considered a cost by most drivers. Such delay not only has a negative value in itself, which may be expressed partially in monetary terms, but it may cause irritation and speeding elsewhere and thus lead to an increased accident rate or even to a contempt for law, a chain of negative consequences that can be difficult to quantify.

An important question, and one of the analyst's important decisions, is the width of the spectrum of consequences to consider. Naming, or listing, the relevant consequences is needed. But which are the relevant ones? We cannot avoid some assessment of the magnitudes and values of the consequences at an early stage. For practical reasons, analysis must be limited: if we consider too many phenomena in the physical, economic, and social environment as being related to the issue under investigation (too many impacts), then the analysis will become expensive, time-consuming, and ineffective. The important consequences are those the decisionmaker will take into account in making his decision, but his list may have to be amplified, for he is also an interested party and may stress beneficial outcomes while neglecting those implying costs or irrelevant to him but detrimental to others.

Therefore, the major responsibility is with the analyst: what consequences to consider is one of the important "secondary decisions" in the study. There is little, if any, theory on which to base this decision. Initial assessments based on experience, common sense, and understanding of the issue are a starting point, but may have to be revised in the course of analysis.

There is one more question related to listing the consequences. How far ahead into the future shall the consequences be considered?

At least two factors influence the answer:

- first, how far-reaching are the objectives (what is the decisionmaker's time horizon), and

- second, how long will the consequences (beneficial and detrimental) last?

These two factors are quite different, and they may be conflicting in the sense that an action taken to achieve a short-term objective may have long-lasting consequences that make it harder to achieve an objective more remote in time. The time horizon of analysis has to be matched to both; the analyst is obliged to tell a short-sighted decisionmaker what the consequences of his action will be in the more distant future.

Predictive models

Analysis predicts what the consequences will be. They cannot be measured or observed; they must be predicted from the present understanding of the future situation and of what the real relations are between the action and its consequences. The process, device, or scheme used for prediction is called a *model.*

The models used in systems analysis may be formal (e.g., mathematical expressions, diagrams, tables), or judgmental (e.g., as formed by the deductions and assessments contained in the mind of an expert). The models most used, on the whole the most useful, and often the only type even considered, are mathematical models. A mathematical model consists of a set of equations and other formal relations that attempt to describe the processes determining the outcome of alternative actions. These models, as do any models, depend for their validity on the quality of the scientific information they represent. A mathematical model is often presented and used in the form of a computer program. Our current capability to design valid and reliable models of these types is limited, particularly for questions of public policy, where social and political considerations tend to dominate. Here, what may be regarded as less satisfactory judgmental models, that depend more, and more directly, on expertise and intuition and are not as precise and manageable, may have to be used to predict the consequences of an alternative. If they are to generate confidence in others, however, they should be made as explicit as possible.

An explicit model of any kind introduces structure and terminology to a problem and provides a means for breaking a complicated decision into smaller tasks that can be handled one at a time. It also serves as an effective means of communication, enabling the participants in a study to make their judgments within a defined context and in proper relation to the judgments of others. Moreover, through feedback—the results of computation in a mathematical model or the criticism of an expert's judgment, for instance—the model can help the analysts and the experts on whom they depend to revise their earlier judgments and thus arrive at a clearer understanding of their subject matter and of the problem.

These secondary characteristics of a model—separating tasks and providing a systematic, efficient, and explicit way to focus judgment and intuition—are crucially important, for they provide a way of tracing out the major consequences when adequate quantitative methods are not available.

It is convenient, in the models, to distinguish two sets of factors that influence simultaneously the outcomes; the consequences y depend on the action a and the state of nature e. "State of nature" is a name given to all exogenous factors, that is, ones beyond control by action a, but which nevertheless influence the consequences y. The important convenience of this approach is that the forecast of the future conditions, and therefore most of the uncertainty, is now contained in the independent, partially random value of e. We can write

$$y = f(a,e), \tag{1}$$

where we mean that "y depends on both a and e," or, stressing the causality, "y is caused by both a and e."

The relation (1) may be considered the general form of a *predictive model*. It is "predictive" in the sense that, given a and e, it determines y. We refer to it as a "model" to stress that, whatever the efforts, our knowledge about the dependencies is restricted and the real relation is different.

We do not imply, by any means, that (1) has some particular form, e.g., that it is a formal mathematical model. It may be a "mental model," contained in the expert's mind, never written down in any form; but nevertheless it may supply statements of the sort: "if action a is taken, given condition e, y will result."

Object system and environment

In analyzing consequences, one always has to make a practical distinction between the *object system* (which we influence) and its *environment*, the *state of nature* or *context*, as we also call it.

From the point of view of the purpose, which is to predict the consequences of a course of action, we may call the environment "the aspects of the world outside the object system that influence the consequences, but are not influenced themselves by the course of action considered."

We take a pragmatic approach in this definition. Rather than stating that everything is related to everything else, we draw a boundary between what has an influence on the consequences that we consider, and what has none (not all the outside world is considered to be the environment), and we draw another boundary between what we influence by an action (the consequences), and what we do not influence.

In other words: although, in principle, the action we take does influence everything else, we choose to divide the relevant world into two parts, the *object system* that we assume to influence, and the *environment*, which is an uninfluenced source of exogenous actions.

The actual decision, that is, setting the boundaries, is subject to decision by the analyst, a decision that may, on the one hand enlighten the analysis and make it a feasible task, but on the other yield superficial, oversimplified, and misleading results. A great deal of clear thinking is necessary to take proper account of the interactions involved, to decide what will be taken into consideration and what can be neglected. Depending on the case, the environment may include, for example, demographic factors, social attitudes, the political and economic situation, and so on.

Model-building limitations; need for experiments

Even in the situations where the phenomena to be predicted are quantifiable, the correctness (validity) of models is limited by several factors: restricted knowledge of the laws of system behavior, inadequate data, inability to deal effectively with very complex relationships, and so on.

The art of model-building has its limitations. Some of these are:

- the data from passive observations alone may not reveal the cause-effect relationships,

- the causal laws that we know, even for physical systems such as chemical reactors, are not enough alone to provide exact models because of the complexity of real systems.

In most cases of model building, both of these sources of information have to be used. They may still be inadequate; then the model-builders may suggest *experiments*. An experiment might consist, for example, of testing a proposed course of action on a sample, and on a parallel control group, observing the results, and then using them both for building a model for arriving at conclusions about the action, and for modifying it before it is applied full-scale.

An experiment can tell how the system reacts in the present, but not how it will react in the future, under changed conditions that cannot be duplicated in the experiment. Because of this principal limitation of experiments, we should not overlook the fact that experimentation alone can hardly be a substitute for other study of a problem, but should rather be considered a complementary activity.

Simulating intervening decisionmakers

The systems that we described so simply by equation (1) may be, internally, very complex. A particular difficulty arises when the system contains one or more decisionmakers whose decisions depend on the action *a*. These decisionmakers who intervene in the process are part of the system; they cannot be considered the environment, because they are influenced by *a*. They can be individuals, social groups, and so on. We have to incorporate them, that is, their behavior, into the model.

For example, we may wish to predict what decrease in gasoline consumption will be caused by a rise in its price; this prediction involves an assumption about the decisions of individual consumers who will decide how much to buy. In the analysis, we can lump all consumers into a single unit (the market), or consider different groups of consumers: farmers, commuters, low, middle, and high income divisions, etc.

It may be reasonable to assume that these other decisionmakers behave rationally; for example, that they maximize their net benefits. Then it would be relatively easy to describe them by some fixed mathematical models. Popular versions of such models are the *demand functions*, which express how much of a commodity the consumer will buy at various prices. These functions are derived by optimizing an assumed objective, the implied objective of the consumer.

In many cases, the assumptions about human decisionmaking cannot be reduced to an optimization problem. We may not know what form of the objective the decisionmakers would optimize. This is a reason for the growing importance—for systems analysis—of the psychological and sociological theory of value and choice. We are unable to predict the consequences of a course of action unless we understand the laws of behavior of the group that will be affected by this action [13].

If we know of the existence, position, and action possibilities of other decisionmakers, we may try to imitate their behavior by appropriately chosen actors. We expect these actors to behave, in the model, in a way that corresponds to what the actual decisionmakers would do in real-world situations, with all the ambiguity and uncertainty. We *simulate* the decisionmakers.

If all of the dependencies, except for human decisionmaking, are programmed into a computer, the whole model becomes an interactive model, or man-machine model, where human decisions interact with input and output data from the computer program.

Models of this kind, although not necessarily involving computers, have been known for a long time under the names of operational games, war games, business games, etc., depending on the context.

Using models

Using a predictive model is in principle very simple: we take the proposed action as an input to it, the assumed or predicted future state of nature as another input, and record the output, that is, the model-predicted consequences.

It is also important to use the model for *sensitivity analysis:* how are the consequences changed if one modifies the action (sensitivity to action) and how are the consequences changed if the exogenous factors change (sensitivity to environmental conditions). A similar investigation, but with respect to major changes in the assumptions about the future state of the world, is sometimes referred to as *contingency analysis*.

In many applications, low sensitivities are important, for the simple reason that no action will ever be implemented with absolute adequacy, no exogenous factor will keep to the forecasted value, and the model from which the results are obtained is never absolutely accurate.

The actual techniques by which the consequences, for given inputs, are predicted depends on the kind of model, for example, whether it is an analytic model (an explicit mathematical relation or formula) or a judgmental (mental) model. However, all kinds of useful models should permit assessing sensitivities.

We are well aware that the future can be determined only in a probabilistic way. It is therefore correct, at least in principle, to ask the model to predict the probabilistic features of the consequences. We may, for example, be interested in the range or interval within which a consequence will be contained with some given (and high) probability. Obtaining answers of this kind requires much information, which will seldom be found in systems-analysis applications. In particular, adequate probabilistic data on the future state of nature, i.e., on future environmental inputs, would have to be available, but seldom are.

We should also mention that the techniques of estimating the probabilistic features of the outcomes may be quite complex and time-consuming. Unless an analytic model is available, a *stochastic computer simulation* can be carried out. In this technique, the computer model is subjected to a large number of suitably generated random inputs, which imitate the stochastic environment. A statistical analysis of the outputs provides the required probabilistic data. This kind of analysis is important in some applications. In many cases, however, a computer simulation is the least desirable model. It is costly, except in the model-building stage, and it has low insight, since it does not show how the observed outcomes are obtained. Nevertheless, it may be the only choice open [14].

In most applications of systems analysis, the scarcity of data and the inaccuracy of models do not permit or justify a precise probabilistic analysis. We should, however, always realize the probabilistic character of the problem and proceed cautiously. A common pitfall, for example, is to take the expected value of the environmental input as a basis for determining the expected value of the outcome. A simple example will explain what happens. Assume a crop increases with humidity, but is more sensitive to drought than to above-average rainfall. Then, calculating the average crop on the basis of average rainfall is wrong, because the losses due to dry years will be more than the gains in the wet years. In more precise terms, what we should do is to calculate the average value of y in equation (1). It cannot be replaced by putting the average value of e into the formula, unless the relation is linear.

Let us focus again on the future aspect of analysis. In the context of using predictive models, we should know e, the state of nature; ahead of time, and this requirement raises two principal questions:

(i) How far ahead *should* we know the state of nature?

(ii) How far ahead *can* we know the state of nature, that is, how far can a reasonable forecast of the state of nature reach (based on past data and available knowledge)?

The answer to the second question is that, whatever the forecasting techniques, the ability to determine the future in terms of reasonable probabilistic confidence is limited. There are, however, many cases in analysis where the future that we must consider is more distant than our forecast of the external conditions can reach. In these cases, the analyst tends to represent the future environment by *scenarios*, i.e., hypothesized chains of events. He is still able to say: if the external events follow scenario No. 1, the results of the action will be ..., but he cannot say much about the probabilities.

A few remarks are appropriate here:

a) As the probabilities of the scenarios are not known, nothing can be said about the *expected* outcome of the action.

b) It is important to consider several scenarios, and to choose them in a systematic way. For example, it is appropriate to consider the scenario that seems to be most likely. However, we may also want to consider other scenarios, structured so as to be more unfavorable to achieving the objective, but which we feel are still likely,

c) An alternative that is very sensitive to a small change in the scenario should be rejected, or redesigned with the purpose of decreasing the sensitivity.

In some applications of systems analysis it is appropriate to replace a probabilistic forecast of the future or an impartial scenario by an active element, an element that will respond to our action in such a way as to purposely upset the potential benefits.

For example, when a plan for developing water resources is being considered, we may ask whether the water demands of all users will be satisfied under all possible circumstances if this plan is implemented. This question calls for an examination of the worst case of the weather and other conditions. We can, for that purpose, treat the state of nature as acting against us. In the model, we can assign the role of nature to an antagonistic player, and thus make another use of *gaming*. Needless to say, to get reasonable conclusions, the action possibilities available to the opponent will have to be bounded in some way, otherwise, no water system could withstand the test. In any case, the game will reveal what exogenous conditions are the most dangerous, and we can then try to assess whether these conditions are likely to happen.

Analyzing infrequent contingencies

In many analyses there is a need to consider *infrequent contingencies*, events or conditions that may happen whose probabilities are low or very low, but which—if they happen—have significant consequences. Usually, these consequences are of a detrimental nature—if they were benefits we would not worry.

The analysis has two parts. The first uses a model or an understanding of the system to determine "what will happen if...;" that is, it looks at various possible failures of system components (a leak in the cooling tubes of a nuclear reactor, for example), or at various possible events in the environment (an earthquake, for example), and it predicts the consequences of the adverse event. Usually, these are serious consequences: a blow-up of the plant, a destruction of the pipeline, etc.

The second part deals with estimating the probability of a given failure or outside event and—when many possible failures or adverse events are considered—estimating the correlations. This analysis means an inquiry into whether the various failures, if they happen, happen independently of each other. This part of the study, i.e., estimating the probabilities, is often referred to as *risk assessment*.

Summary remarks

Let us come back to the main questions addressed at the beginning of this section:

- What will the future of the world be if the action is taken?

- How certain is the answer that analysis can supply?

In many worthwhile applications of systems analysis, in spite of all the efforts that can be put into model-building and forecasting activities, we usually cannot claim that the consequences we predict will happen with reasonably high probabilities. It should be understood from what has been said that this uncertainty in the answer cannot be entirely overcome. Is it, then, reasonable to spend money and time on systems analysis, to build and use models in cases where they cannot predict accurately?

The answer is yes. For one thing, the decisionmaker has to make a decision anyway, and even imperfect assistance by analysis may be better than pure judgment and intuition. Second, analysis may permit comparing alternatives, even if the absolute accuracy of predicting the consequences is low.

For example, assume there is no probabilistic forecast of the future, but only a few scenarios. If we then detect, by a consistent model-based analysis, that the consequences of action a1 are better than those of action a2 under all, or most, of the representative scenarios, this result is a useful indication.

Other useful indications that a model can provide are indications of sensitivity: a course of action that makes a system insensitive to exogenous factors, or makes it able to recover *(resilient)*, is a preferable one, even if we do not know its consequences exactly.

4.5 FORECASTING PROBABLE STATES OF THE WORLD

Forecasting in systems analysis

Forecasting is needed in every systems-analysis study. Before any action can be evaluated, we require a forecast of the future "state of the world," as a context in which the consequences of an action can be predicted and compared. Forecasting is needed even when we just want to discover if some action is needed. Weather forecasting is one example, econometric forecasts used to draw inferences about the future state of national economies are another. We should note that, although sophisticated models and extensive statistical data analysis are used in these two forecasts, we do not insist on knowing the cause-effect relationships. The forecasting models show *correlations*, but may fail to show *dependencies*. It is a common pitfall to neglect the difference, and thus to draw false conclusions about what a deliberate action may bring about. For example, we cannot cause rainfall by forcing the birds to fly at low altitudes, although the two facts are known to be strongly correlated in some climates (because both of them are effects of the same cause—air humidity).

A forecast of the future state of the world is, of course, also needed when one wants to know what the probable consequences of an intended action are. As pointed out in section 4.4, it is important, in this case, to distinguish between the phenomena or features of the world that will be influenced by the action (the object system), and the ones that will not (the environment, state of nature).

If our confidence in the accuracy of the forecast is not extremely high, we may want to carry out the analysis for several different projections of possible states of the world.

Forecasting techniques

Forecasting can be done in a variety of ways. Forecasting techniques range from expert judgements to mathematical forecasting models. Whatever technique is being used, a forecast is always based on past and current data, observations, or measurements. When based on expert judgment, it is done to a large extent implicitly, and in model-based forecasting, explicitly.

It may be appropriate, at this point, to indicate that even the best forecasting technique determines the future only in a probabilistic way. For example, it may—in the best case—state the expected value and the variance, or the confidence interval within which the value will be contained with some probability. The variance, or the confidence interval just mentioned, is bound to increase as the future considered is more distant. In view of the available data, a forecasting technique should be chosen that is not too sophisticated for the circumstances. In many cases simple forecasting models are often as good as the very complex ones, if data are scarce. It may also be impractical, in the early stages of analysis when more qualitative answers are sought, to use the more complex forecasting models. Simplified versions may be useful at these stages.

Data and results

In any forecasting technique, we make the essential assumption that the future is partially determined by the past, on which data can be made available. This assumption implies these important questions related to the *data* needed for a reliable forecast:

(i) How far into the past should the record reach?

(ii) How broad should the observations be, i.e., how many different phenomena must be observed to forecast one selected phenomenon in the future?

(iii) Can we trade the length of record for broadness of observations, and, if so, to what extent?

(iv) How far ahead can we infer from the data available?

We stress that one should not overestimate the power and possibilities of forecasting techniques based on statistical data and formal models. For one thing, the data may not be rich enough to provide the necessary length and broadness of the record. Secondly, the phenomena in the past were observed (measured) with errors. Thirdly, there are phenomena to be forecast in some systems analyses that are related to phenomena in the past that are either not measurable, or missing from the statistics. For example, long-term forecasts of changes in technology due to inventions, and forecasts of changes in societal and political attitudes, have natures that deny success to formal model-based forecasting methods.

There certainly are many other cases where expert-based, *judgmental forecasting* may be appropriate, because human experience and intuition may—implicitly and even unconsciously—make use of correlations and associations that cannot readily be formalized. There are various methods for organizing and assessing the results of expert evaluation; they are superior to the usual committee activity for using groups of experts for forecasting, parameter estimation, the ranking of alternatives, and certain other purposes.

4.6 COMPARING AND RANKING ALTERNATIVES

Difficulties of comparison

Assume the alternatives have been selected and screened, and the presumed consequences of each determined. How can we compare them? An obvious method is to display the alternatives in a suitable framework so that the differences and similarities stand out. The analyst may also do more; for example, he may rank the alternatives according to one or more specified criteria, so that the decisionmaker's choice is made easier.

There are several reasons why ranking alternatives is difficult, except in the simplest cases.

- In most practical cases alternative A may be superior to B in some aspects and inferior in the others.

- The diverse consequences of an alternative cannot be aggregated into a single performance index, which bears a satisfactory relation to attaining the objectives.

- When outcomes are spread over time, and not in the same way for various alternatives, their rankings may change with time.

- There may exist consequences that are nonquantifiable on a generally accepted basis and that may be quantified by judgment only.

- The future conditions under which a proposed alternative will have to function are unpredictable. At the same time, the range of probable future conditions is wide and bears strongly on the presumed consequences.

Nevertheless, in spite of the difficulties and all the incommensurability of various effects and consequences of the alternatives, a choice has to be made by the decisionmaker.

We are concerned about the extent to which this decision can be assisted by the analyst, for example, the extent and the means by which we can reduce the variety of features of each alternative into possibly few, but nevertheless reliable, indicators.

A danger is oversimplification, i.e., of trying to merge too many things into a single index value. One should not neglect the fact that a subjective judgement by the decisionmaker on a set of displayed impacts may be more adequate than an index arrived at by arbitrary quantification, questionable arguments, and value estimates by the analysts. Simple judgement may quite often lead to the right decision, as opposed to the theory-based decision done in the right way. There is, however, a lot of significant research devoted to the problem of analyzing and modeling cognitive value systems.

Judgmental comparison and ranking

The simplest method, as mentioned above, is to display the impacts of the alternatives to the decisionmaker. Such a display is sometimes referred to as a scorecard (see the example in Chapter 8). A scorecard aims to present the decisionmaker with the full spectrum of consequences, both good and bad, with an indication of who gets the benefits and who pays the costs. The decisionmaker can superimpose on this relatively objective information his feelings for the values, as well as incorporate the value judgments of the society he represents. This approach makes it also possible to show sensitivities, that is, to show how the impacts change when parameters and external conditions vary. One merely prepares a scorecard for the same alternatives under the changed conditions, and superimposes it on the previous one.

The scorecard is also effective for multiple decisionmakers, for each individual may form his own opinion based on his preferences and prejudices and a consensus can then be worked out through committee action. It is also easily understood by the public at large.

Any evaluation and ranking of the alternatives by analysts or experts may ignore important factors known to the decisionmaker but never made explicit to the analysts and experts. Therefore, such ranking may be unsatisfactory to the decisionmaker. For this reason alone, he should always be presented with the major alternatives and their impacts. In other words, when we present a decisionmaker with the result of someone else's evaluation, we should produce the scorecard for all highly ranked alternatives.

The sheer mass of information, however, makes the use of indices of various sorts attractive.

Cost-effectiveness and cost-benefit criteria

For the purpose of comparing and ranking alternatives, one often tries to describe their relative merits by means of one, or at most a few, indicators (index value, figure of merit, or objective function). Any such approach has to sacrifice the details, the individual features of the alternatives, for the sake of making comparison easier.

Cost-effectiveness can be used to rank alternatives when there is a single dominant objective and the effectiveness of the various alternatives in attaining this objective can be measured on a single scale that is directly related to the objective, or is a good proxy for it. Alternatives are ranked either in terms of decreasing effectiveness for equal cost or, less frequently, in terms of increasing cost for equal effectiveness. Sometimes the ratio of cost to effectiveness is used, but this practice is open to all the objections that apply to the use of ratios as criteria, for example, because this kind of criteria mask the differences in scale [15].

The cost-effectiveness criterion is open to a number of objections. For one, even in the simplest cases, effectiveness may not measure value, which depends on the particular decisionmaker. For another, if the ranking is close, the decisionmaker may want to consider secondary effects.

Another objection is that cost as used in cost-effectiveness reflects only the costs that are inputs—the money, resources, time, and manpower required to implement and maintain an alternative. The penalties or losses that may accompany an implemented alternative—it may, for instance, interfere with something else that is wanted or bring undesirable consequences to other people—are costs that are not taken into account.

Finally, even if cost and effectiveness are properly determined, the decisionmaker is still faced with the problem of what to choose. He needs some way to set the scale of effort—either a cost he must not exceed or an effectiveness level he needs to achieve. The ratio of cost to effectiveness may not be a satisfactory guide, even if he is totally uninterested in the scale of effort.

Cost-benefit analysis can, in a theoretical sense, handle the difficulties associated with cost-effectiveness. In this approach, the costs and benefits that follow each choice of an alternative, properly associated with the times at which they occur, are measured in the same units, usually monetary. Whether such a transfer into monetary terms can be viewed as valid is, however, a difficult question.

Value and utility approaches.

A decisionmaker for whom a systems analysis is made faces the following situation: there are several alternatives from which to choose, each of them with many consequences (impacts), and, moreover, the impacts predicted for each alternative are different under different future conditions. One does not know which of the future conditions (referred to as states of nature, section 4.4) will occur.

If a scorecard approach is used, there will be as many scorecards for each alternative as there are scenarios of the future to be considered. On each scorecard the number of entries is the number of impacts multiplied by the number of alternatives to be evaluated and compared. The result may be that the amount of data is far too large for the decisionmaker, without some aggregation, to make a judgmental ranking and choice.

It is therefore understandable that there is a tendency to evaluate each alternative by a single indicator such as *effectiveness for fixed cost* or *net benefit* in a cost-benefit analysis. We still have, of course, the various possible futures, and hence several different values of the chosen indicator for each alternative. Nevertheless, the display of alternatives is more lucid and transparent.

The concept of using a single indicator for many noncommensurable features of an object, in our case, an assumed alternative, is well known to decision analysts, who have formalized it as a *multiattribute value function*. In this approach one tries to build a function by which a value v is assigned to the consequence of an alternative, whereby this consequence is assmed to have n different value-relevant attributes:

$$v = f(y_1, y_2 \cdots, y_n), \tag{2}$$

where $y_1, y_2 \cdots, y_n$ are the value-relevant attributes measured on their appropriate scales.

The function in (2) is a model of the decisionmaker's value system. It has to be established on the basis of his preferences, that is, of his individual judgments, and this is where the difficulties arise. In practice, it is, for many reasons, hard to obtain a value function that could replace the actual decision on a complex and unique issue. It is possible, however, for multiattribute value functions to be used as a guide or directive in the initial selection, design, and fine tuning of alternatives, or as one of the ranking criteria to be compared with rankings done by other means. A public official's preferences are, in general, the preferences of the people he represents; it is through this association that the analyst can get an idea of the decisionmaker's preferences. One still faces the problem of uncertainty: even if we agree to evaluate the alternatives by a single indicator for each of the possible states of nature, how should they be ranked, since we do not know which of these states of nature will occur?

Let us assume that, from one source or another, the probabilities of the various future states of nature are known or can be estimated. It seems quite natural in this case, or at least simplest, to rank the alternatives on the basis of the mathematical expectation (expected value) of the outcome.

Using the multiattribute value function expressed by equation (2), which assigns a single value indicator to a given alternative under one state of nature, one can calculate the average value for each alternative over all possible states of nature. Then, the alternatives can be ranked according to the average, i.e., expected, values.

It should be noted, however, that a straightforward average may not indicate the choice that a given decisionmaker would make.

To take account of this, we use the notion of *utility*, a basic concept used in the theory of *decision under risk*. This theory assigns utilities to consequences in such a way that ranking expected utilities of alternatives is the same as the decisionmaker's preference order for the same alternatives.

Utilities are assigned to consequences by means of *utility functions;* a utility function describes the attitude of the given decisionmaker towards risk, and is thus different for risk-averse and risk-prone decisionmakers.

Direct use of utility theory, i.e., of utility functions and the expected utility principle, for ranking alternatives and, in particular, for a final choice, cannot be recommended without reservation. Assigning utility via utility functions involves a great deal of judgment by the analyst; several simplifying assumptions with respect to the form of these functions are also indispensable. Nevertheless, as in the case of miltiattribute value functions, expected utility may be valuable as one of the means by which the alternatives can be screened and assigned a tentative ranking, even if it cannot be recommended as a unique and ultimate criterion for choice.

Summary remarks

Relatively little can be added to what was said in the first paragraph of this section: comparing alternatives is, in all practical cases, difficult. We should also remember that, although comparison and choice go together, the two parts are done by different people. It is the duty of the analyst to provide a comparison of alternatives and possibly a ranking, but it is the right and responsibility of the decisionmaker to make the choice.

It is, therefore, reasonable not to rely entirely on the rankings provided by cost-benefit, multiattribute value functions, or utility functions. A scorecard of the alternatives, reduced, perhaps, to the most relevant attributes, should accompany any rank-ordered list of alternatives.

The analyst should not be upset if the choice of the decisionmaker is the third-ranked or fourth-ranked alternative. Such a choice indicates only that there are additional aspects and values that the decisionmaker did not disclose before, or that were misunderstood by the analyst. The analysis at this stage may be considered a success if the decisionmaker has made an analysis-based decision in the sense that he has chosen a course of action taking into account consequences that have been duly and appropriately analyzed. We must remember, however, that the analyst's goal is not merely to find the course of action best suited to achieve the decisionmaker's objectives and satisfy his constraints, but to find the course of action closest to this ideal that can be accepted by the other participants in the decisionmaking process and then implemented without undesirable modification, or extra cost and delay.

As mentioned in section 4.1 the analyst's role does not necessarily end at the choice by the decisionmakers of a particular course of action. Analysts will usually be called on to assist with implementation, especially in the early part of this process when there may be a need to interpret aspects of the program, as well as for modifications due to circumstances that were impossible to anticipate earlier. Analysts, although probably not the same analysts, will also play an important role when it comes to evaluating the results of the implemented action, and the original analysis itself.

REFERENCES

1. Allison, Graham T. (1971) *Essence of Decision,* Boston: Little Brown, and Company, p. 10.

2. Rein, Martin, and Sheldon H. White, (1977) "Policy Research: Belief and Doubt," *Policy Analysis* 3:2 (Spring), p. 249.

3. Lynn, Laurence E., Jr. (1978) "The Question of Relevance," in L.E. Lynn, Jr., ed., *Knowledge and Policy: The Uncertain Connection,* Washington, D.C.: National Research Council.

4. Nelson, Richard R. (1974) "Intellectualizing About the Moon-Ghetto Metaphor: A Study of the Current Malaise of Rational Analysis of Social Problems," *Policy Sciences* 5, p. 381.

5. Quade, E.S. (1968) "Introduction", in E.S. Quade and W.I. Boucher, eds., *Systems Analysis and Policy Planning,* New York: American Elsevier.

6. Quade, E.S. (1975) *Analysis for Public Decisions* New York: American Elsevier, p. 32-67.

7. Goeller, Bruce, F. (1969) "Methodology for Determining Traffic Safety Priorities: A Collision Prediction Model," P-3962, The Rand Corporation, Santa Monica, California.

8. White, D.J. (1975) *Decision Methodology*, London: Wiley.

9. Keeney, R.L., and Howard Raiffa (1977) *Decision with Multiple Objectives: Preferences and Value Tradeoffs*, New York: Wiley.

10. Bell, David E., R.L. Keeney, and Howard Raiffa, editors (1977) *Conflicting Objectives in Decisions*, Chichester, England: Wiley.

11. Simon, H. (1964) "On the Concept of Organizational Goal", *Administrative Science Quarterly*, Vol. 9, No. 1.

12. Majone, G., and E.S. Quade (1979) *Pitfalls of Analysis* Chichester, England: Wiley.

13. Stokey, Elizabeth, and Richard Zeckhauser (1978) *A Primer for Policy Analysts*, New York: Norton.

14. Chapter 7 of this Handbook.

15. Hitch, C.J., and R. McKean (1960) *The Economics of Defense in the Nuclear Age*, Cambridge, Mass.: Harvard University Press.