(2)

Onward Through the Fog

Uncertainty and Management Adaptation in Systems Analysis and Design

James S. Hodges

RAND | PROJECT AIR FORCE
ARROYO CENTER
NATIONAL DEFENSE RESEARCH INSTITUTE

The research described in this report was sponsored by the United States Air Force, Contract No. F49620-86-C-0008; by the United States Army, Contract No. MDA903-86-C-0059; and by the Office of the Secretary of Defense (OSD), under RAND's National Defense Research Institute, an OSD-supported federally funded research and development center, Contract No. MDA903-85-C-0030.

ISBN: 0-8330-1073-5

The RAND Publication Series: The Report is the principal publication documenting and transmitting RAND's major research findings and final research results. The RAND Note reports other outputs of sponsored research for general distribution. Publications of The RAND Corporation do not necessarily reflect the opinions or policies of the sponsors of RAND research.

Published by The RAND Corporation
1700 Main Street, P.O. Box 2138, Santa Monica, CA 90406-2138

R-3760-AF/A/OSD

Onward Through the Fog

Uncertainty and Management Adaptation in Systems Analysis and Design

James S. Hodges
with Raymond A. Pyles

July 1990

Prepared for the
United States Air Force
United States Army
Office of the Secretary of Defense

RAND

PREFACE

Policy analysts have always lived with great uncertainty. We have long had tools for handling some of that uncertainty. But now that policy analysis has been around for a while, work in many fields has fallen into stereotyped problem formulations and analytical approaches. In particular, treatments of uncertainty are typically quite incomplete and often conceptually wrong. This report argues that these shortcomings produce pervasive systematic biases in analyses. Problem formulations suppress uncertainty about the world and ignore the capacity of the analyzed systems to adapt to the unanticipated. System design ignores opportunities to hedge or respond instead of predicting. System evaluation discounts strategies that hedge or respond.

Policy analyses change in specific ways if their problems are formulated with a fuller recognition of uncertainty. Management capability—the ability to respond and adapt—becomes a resource, with a cost and an effectiveness, to which funds can be allocated in competition with more familiar uses of funds. Analytical tools change to incorporate options for management. Given that uncertainty is pervasive and generally cannot be counted completely, the logic that connects the analysis to the world changes as well: Our models are not necessarily adequate representations of reality.

Although its main example is an Air Force logistics problem, this report is intended for policy analysts in all fields. It should also be of interest to operations researchers, statisticians, and researchers in other quantitative fields.

This research drew support from RAND's three federally funded research and development centers:

- Arroyo Center
- National Defense Research Institute
- Project AIR FORCE

The work was performed under the project entitled "Logistics Monographs."

SUMMARY

If events were predictable, policy analysis would be simple. For example, if Air Force planners knew which parts would fail, and where and when they would fail, it would be straightforward to schedule repair and procurement of spares and repair capacity. But planners do not know what will fail or where or when and must plan accordingly. Every policy analytic problem shares this uncertainty to some degree. The common mode of analysis omits crucial sources of uncertainty and overlooks the ability of systems to respond to the unexpected. These two omissions systematically bias design and evaluation of systems.

The process of analysis is broken into five stages: (1) setting the context, (2) choosing alternatives, (3) predicting consequences, (4) valuing outcomes, and (5) making a choice. This critique of the common approach applies mostly to the first four stages, which rely on a three-part representation of the actual problem:

- a *tasking* posited for the system under study, consisting of responsibilities assigned to the system and conditions under which they will be discharged,
- a *mechanism* that turns the tasking into a load on the system, and
- the system under study, which is taxed by the load and consumes *and manages* resources to service it. For emphasis, we call it the *managed system*.

Typically, policymakers and analysts manipulate resources supplied to the managed system. For example, in Air Force spares and repair requirements calculations, one tasking—a hypothetical war, part failure rates, and repair and resupply times—is derived from higher-level guidance. The mechanism consists mostly of models that relate flying hours to numbers of parts needing repair. The managed system—the part repair and resupply system—is modeled without constraints on repair or transportation capacity. The requirement computation determines how much repair capacity and stock to buy for the managed system without regard for the way the repair capacity and stock are managed.

The common mode of analysis omits sources of uncertainty from each part of the representation. One such source is *statistical uncertainty:* variability observed in repeatable phenomena. Another source of uncertainty arises in phenomena that are not repeatable, not observed, or both: *state-of-the-world* uncertainty. It comes in three main types: extrapolation away from observed conditions and behavior of people *inside* or *outside* the mechanism or managed system.

In Air Force spare parts and repair requirement calculations, some variation in part failures is represented, but substantially less than is observed in peacetime. State-of-the-world uncertainty is ignored. This latter includes uncertainty about:

- The tasking: Which war will be fought, under what conditions?
- The mechanism: Can peacetime failure behavior be extrapolated to wartime?
- The managed system: Adaptations made by the repair and resupply system are ignored; what capability will they provide in wartime, and at what cost?

This example is typical of nonmilitary analyses as well as of military analyses.

If things are so bad, why aren't more systems hopelessly bogged down? Because mechanisms and managed systems are not inert, as they are modeled, but adaptive. In the Air Force example, more and less informal adaptations in the repair and resupply system can shift resources toward critical parts. Adaptive behavior provides a capability to function in the face of uncertainty, and analysts may help managers more by devising (and evaluating) ways for them to adapt, instead of by trying to help them predict the unpredictable.

This approach has been explored to some extent with a study of some Air Force spare parts initiatives, called CLOUT (Coupling Logistics to Operations to meet Uncertainty and the Threat). This study added to the usual analyses:

- Uncertainty, in the form of actual peacetime variability in part failures and base attack damage;
- Management adaptations, such as sharing of repair resources among bases.

The CLOUT analysis indicated that these measures could produce a substantial payoff in aircraft availability in combat theaters.

The critique here has implications for the common mode of policy analysis, mostly about design and evaluation of systems, which comprise the first four of the five stages of policy analysis. The implications include:

- System Design (Stage 2, Choosing Alternatives): Systems need strategies for coping with uncertainty. We identify several strategies as a source of ideas for specific design problems and to illustrate the importance of understanding the relative value of management (adaptiveness) and goods (ampleness).
- System Evaluation (Stage 3, Predicting Consequences, and Stage 4, Valuing Outcomes): Models can represent more statistical uncertainty, and analyses can allow for more state-of-the world uncertainty in a sensitivity analysis by considering variants on the tasking. But often these two methods are not enough to connect the models and sensitivity analysis to the world: They can capture more uncertainty, but not enough. In such cases, other kinds of arguments are needed to make a clean logical connection between the analysis and the world. We examine two: prototyping (or similar operational tests) and a fortiori arguments. Prototyping is a superb method but is available for only a limited class of problems. For other situations, the a fortiori argument may be available. The CLOUT analyses are an example. The model used in those analyses treats part failures as more predictable than they really are, and allows fewer management adaptations than will probably be available. In this world, the CLOUT initiatives are preferable to current nonadaptive methods. In an actual war, conditions will probably be more favorable to CLOUT, so a fortiori, CLOUT appears to prevail.
- Setting the Tasking (Stage 1, Setting the Context): Analyses must reflect uncertainty about the tasking by considering a range of taskings. This does not mean, for example, that budget requests must be ranges instead of single numbers; it does mean that single budget or requirement numbers must be evaluated against a range of taskings instead of a single tasking. We conjecture that it will be fruitful to explore choosing taskings to drive a fortiori arguments like the above.

ACKNOWLEDGMENTS

These RAND colleagues generously read this report or a previous version and offered comments and suggestions: Jack Abell, Carl Builder, Steven Carroll, Gordon Crawford, Gene Fisher, Lionel Galway, Bruce Goeller, James Hammitt, David Kassing, Robert Levine, Thomas Lippiatt, Lynn Ordway, John Rolph, Jack Stockfisch, and Warren Walker. I. K. Cohen instigated this work and provided indispensable intellectual guidance and prodding throughout. Paul Steinberg helped make the report more user-friendly. Any remaining errors of fact or interpretation are the responsibility of the authors.

CONTENTS

I. INTRODUCTION

BACKGROUND

If events were predictable, policy analysis would be simple. For example, if Air Force planners knew which aircraft parts would fail and where and when they would fail, it would be straightforward to schedule repair of parts, to buy parts and repair capability, and to make budgets. But nobody can accurately predict which parts will fail, when they will fail, or where they will fail, even in peacetime. The wars for which we prepare are also shrouded in uncertainty. How much will our aircraft be called on to fly, and on what kinds of missions? How much damage will our repair assets sustain? Will parts fail in wartime as they do in peacetime? Again, no one knows.

The Air Force captures some of this uncertainty in its computations and analyses about spare parts and repair. For example, uncertainty about part failures is described by simple probability distributions with a simple extrapolation to wartime. However, much uncertainty is omitted from all the Air Force's computations and analyses. For example, apart from the simple extrapolation just mentioned, all wartime uncertainty is omitted from computations used for budgets, requirements estimates, and operating the spares and repair system.

Not coincidentally, these same computations generally use a rudimentary model to represent the system that manages spare parts and repair, if they represent it at all. Thus, repair and transportation capacity are usually assumed to be infinite, and repairs are assumed to be scheduled based on a first-come-first-served rule. As a result, the system's capacity to adapt and respond to the unexpected is simplified away, and the implications of that capacity are lost.

OBJECTIVE

The common mode of analysis, as illustrated by the Air Force example just mentioned, omits crucial sources of uncertainty and overlooks the ability of systems to respond to the unexpected. These flaws introduce a systematic bias into analyses and, thus, into the systems that are motivated and evaluated by those analyses. Analyses serve many purposes: Some are done repeatedly to drive operations; others are done once, such as for purchases; and others are performed for still different purposes. The omissions play out differently in the different roles that analyses fill, and our critique has implications for all these roles. Further, although our main example will be the case of Air Force spare parts and repair, the report's implications apply to all areas in which policy analysis is done. We indicate this with brief examples from munitions requirements computations, disposal of high-level radioactive waste, and fire department management, among others.

ORGANIZATION OF THE REPORT

Section II describes the common mode of policy analysis and identifies the two main shortcomings raised in this section. The next two sections discuss them in turn. Section III categorizes some varieties of uncertainty relevant to policy analysts and gives some examples of how they are commonly represented. The reader is then asked to recall that our military organizations are not hopelessly bogged down, and Sec. IV argues that this is so because managers are not as inert as our models of them: Managers respond to unpredicted circumstances. Their responses provide system capability, the costs and benefits of which are generally unknown and should be evaluated. Further, these responses might be supplemented by other means for responding that are profitable relative to competing uses of resources.

Designing and evaluating systems are the subject of Sec. V. The fundamental design idea that follows from Secs. II through IV is that management is a resource with costs and benefits, which is most likely to be useful in situations of great uncertainty. We present a collection of generic strategies for uncertain situations, mostly to illustrate the tradeoff between buying management and buying other things, but also as a source of design ideas. For system evaluation, we discuss such familiar solutions as adding some statistical variability and management adaptations to models or performing more thorough sensitivity analyses. We do not denigrate the usefulness of these methods when they are appropriate, but we argue that they are sufficient far less often than standard treatments suggest. When they are insufficient, which we believe to be the case most of the time, analyses based on models and sensitivity analysis need a different kind of logic to connect them to the world. We discuss two logically clear alternatives: implementing a proposed system as a prototype and employing a fortiori arguments.

II. THE COMMON MODE OF POLICY ANALYSIS AND ITS SHORTCOMINGS

SUMMARY OF THE COMMON MODE[1]

Stokey and Zeckhauser (1978, p. 5) give a typical breakdown of policy analysis into five stages: (1) establishing the context, (2) laying out the alternatives, (3) predicting the consequences, (4) valuing the outcomes, and (5) making a choice. This report is mostly a critique of the common approach to the first four stages, although it has implications for stage 5. Stage 2 corresponds to what we have called system design, and stages 3 and 4 together constitute what we have called evaluation. Stage 1, establishing the context, and stages 3 and 4, evaluation, generally use a stylized representation of the problem.

The stylized representation may have as many as three parts: (1) a *tasking*[2] posited for the system under study, (2) a *mechanism* that turns the tasking into a load on the system, and (3) the *managed system* under study, which is taxed by the load and consumes resources to service it. Policymakers and analysts manipulate resources supplied to the system in the hope of achieving some desirable outcome.

Tasking

A system cannot be asked to do everything that might need to be done in any conceivable circumstance. It must be given specific *responsibilities* and *conditions* in which it is expected to meet those responsibilities. These constitute its *tasking*[3] and are the main burden of stage 1, Setting the Context. *Responsibilities* are chosen generally by the highest-level policymakers, but operators and analysts within the mechanism and the managed system make many detailed choices. *Conditions* are features of the world with which, in the common mode of analysis, taskers and operators are told they must live. The condi-

[1]It is not our object to do the definitive survey on policy analysis or uncertainty but to present a point of view and enough references that an interested reader can pursue aspects of it.

[2]This term is sometimes used for very detailed instructions given to operators. We use it for high-level assignments given to organizations.

[3]Terminology varies tremendously in this area. Draper et al. (1987) follow one common usage by lumping both responsibilities and conditions under "scenarios." We prefer to distinguish the two.

tions are treated as if they are either truly uncontrollable or not within the purview of those setting the tasking.

In the context of the Air Force example we are using, Air Force planning agencies derive a detailed tasking from guidance provided by higher authorities, to compute spare parts investments and repair requirements. The responsibilities in the tasking consist of the war that the Air Force will try to be ready to fight: a single deployment schedule, flying program, and support plan. The chosen war is, of course, constrained by the possibilities allowed by the current force structure. The conditions in the tasking are a single set of part failure rates, attrition rates for aircraft in combat, times required to repair and ship parts, and many similar things. These are treated as if they are unchangeable and are used to compute spare part purchase and repair requirements as well as war reserve material requirements.[4] We will argue later that some of these conditions are in fact subject to the control of managers and sometimes should be treated as such.

In summary, then, in the stylized representation used in the common mode of analysis, the tasking is represented by one responsibility or a small number of them and by conditions under which the system will face those responsibilities.

Mechanism

The tasking itself does not produce a load on the system under study, but it does provide inputs to a *mechanism* that produces the load. In the Air Force's spares and repair requirement calculations, the tasking is turned into demands on the repair and resupply system by models that relate flying hours or sorties to probability distributions of demands for replacement parts. In general, the mechanism in the stylized representation is a model that takes as inputs the responsibilities and conditions of the tasking and produces a load on the system under study.

[4]When multiple wartime operational taskings exist for a unit, the analysts assign it a "worst case" support plan, but they do not evaluate the consequences of potential deviations in other factors, such as flying hours or repair times, that might affect requirements.

Managed System

Once the mechanism has produced a load on the system, that system consumes resources to service the load. In fact, the system does not simply draw in new resources and consume them, it also manages resources that it owns. We use the term *managed system* to emphasize this.[5]

In the Air Force example, the repair and resupply system removes broken parts from planes, replaces them or orders replacements, and junks or repairs the broken parts. Most models of this process treat the material flows as if they move immutably through an inert and infinitely expandable pipe. The models ignore the effects on system performance of limited transportation and repair and of adaptive behavior within the system. In some analyses, the managed system is not explicitly represented, even though it may be of central importance. For example, the Army's munitions requirements computation[6] does not model supply at all. It assumes that the supply system delivers the right munitions to the right place at the right time.

The notions of tasking, mechanism, and managed system relate straightforwardly to the five stages of policy analysis. Stage 1, establishing the context, is mostly about specifying the tasking. Stage 2, laying out the alternatives, most naturally involves the managed system, although it may involve constraints placed on elements of the mechanism. Stage 3, predicting the consequences, involves the behavior of both the mechanism and the managed system. Stage 4, valuing the outcomes, is generally an evaluation of the managed system and of the adequacy of the resources supplied to it.

THREE ROLES OF POLICY ANALYSIS

Builder (1989, Ch. 9) describes several roles policy analyses may fill. He makes particular mention of three roles: operations analysis, systems analysis, and requirements analysis. Operations analysis is the analysis of an existing process or thing that can be measured. Systems analysis is the comparison of alternative things or processes that are often mostly hypothetical and thus cannot be measured. Requirements analysis takes place after a systems analysis has se-

[5]In the following discussion, the term "manager" refers to any person in the managed system who can react effectively to events. In our lexicon, a supply clerk who expedites badly needed requisitions is as much a manager as a shop chief who sets repair priorities. Managers are those who adapt to changing circumstances by reallocating time, effort, material, money, or information.

[6]See Girardini (1989).

lected an alternative and results in a specific design or a decision about how much of the thing to buy. In the Air Force example, a systems analysis would compare alternative ways of organizing the spares and repair system, a requirements analysis would determine how much inventory and repair capacity to buy, and an operations analysis would examine the chosen system in operation and adjust the earlier choices. All five stages of a policy analysis and each of the notions of tasking, mechanism, and managed system may be present in any of these kinds of analyses, although some may be implicit or vestigial.

TWO SHORTCOMINGS OF THE COMMON MODE

Most analyses proceed as if some things are fixed, known characteristics of the world; other things not fixed by nature can be set by assumption; and everything else can be chosen by policymakers or analysts working for policymakers. In practice, this means that the mechanism and the managed system are usually treated as fixed, known characteristics of the world. All too often, so is the tasking. The resources available to the managed system to meet the predicted load remain to be set by policymakers or their analysts. Thus, spares and repair analysts add more components or repair authorizations to improve aircraft availability or the chance that an item will be in stock under a single peacetime or wartime tasking. These analyses usually attempt to discover the single optimum resource mix needed to meet the assigned tasking, either exactly or with some confidence level. They generally make no certification about whether that resource mix would be appropriate for other taskings.

This common approach has two important shortcomings. First, large and often dominant sources of uncertainty are omitted. Although it is traditional to use probability distributions to represent some uncertainty, it is also traditional to assume away sources of uncertainty that may dwarf those that are represented. For example, the responsibilities and conditions in the tasking can rarely be predicted with confidence, models of the mechanism and managed system are not often validated, and when they are validated they usually show discouraging inaccuracy. All these facts diminish the confidence of the predictions embodied in the tasking, mechanism, and managed system and, if properly counted, will change the outcomes of analyses.

The common approach usually treats mechanism and managed system as if they are inert, ignoring an important source of uncertainty in predictions and possibilities for designing better systems.

As an example of the first point, models used to forecast spare parts requirements assume that the behavior of pilots, ground crews, and repair managers do not affect demands. Conversely, they assume that the volume of demands or the needs of the supported force do not affect the performance of the repair and supply system. These models of complex human systems ignore the interconnections among systems and the effects of incentives and exigency on the people in those systems. As for the second point, by treating the managed system as inert, analysts may forfeit the most important messages in an analysis: Managers must and do respond to the unexpected, the system should provide "levers" with which they can respond, and analysts must evaluate the costs and benefits of such proposed levers as alternatives to other uses of resources.

III. UNCERTAINTY IS PERVASIVE BUT REPRESENTED INADEQUATELY

TWO KINDS OF UNCERTAINTY

Statistical Uncertainty[1]

Statistical uncertainty is variability *observed* in *repeatable* phenomena. "Observed" means that one must have some measurements of the phenomena. A phenomenon is "repeatable" if the conditions are, or can be, created repeatedly and if the entities subjected to the conditions are either the same each time, or are different but indistinguishable. (A technical condition called "exchangeability" is required, but the informal terms "repeatable" and "indistinguishable" are adequate for our purpose.) This implies a stability or persistence of the variation through time, in aggregate terms. When some phenomenon is observable and a case can be made that it is repeatable, it is straightforward to reach consensus on how to describe the uncertainty of future observations of it. In this sense, it is uncontroversial to speak of the existence of a probability distribution from which future observations of the phenomenon will be drawn.

Taken at face value, the notion of statistical uncertainty postulates that invisible random number generators are out there in the world producing variability. This postulate does not withstand hard scrutiny: We use probabilistic models when we cannot or choose not to model in greater detail[2] or when we deliberately introduce randomness, as in random survey samples. Nonetheless, the idea can be

[1]The notion of "statistical uncertainty," as distinct from state-of-the-world uncertainty, goes back at least to Frank Knight in the 1920s. The term "statistical uncertainty" was used by Fisher (1971) and Quade (1982), but several other names have been used. Classical statistical theory calls it "probability," although nowadays that term generally describes other kinds of uncertainty as well. Steinbrunner (1974) called it "risk," and Hitch (1960) called it "insurable risk," in the sense that insurance companies can study and treat large classes of individuals with given characteristics as the bomb tests are treated in the text. Madansky (1968) called it "objective probability," which captures the idea that the variability is inherent in the measured phenomenon, not the observer, and can be observed and perhaps definitively characterized.

[2]Two canonical examples of random mechanisms are coin flips and subatomic particles. Coin flips are deterministic given initial velocity and revolutions per second (DeGroot, 1986). We find writing on subatomic particles to be either mistaken or obscure on the subject of inherent randomness, although it is most unlikely that current probabilistic models of subatomic particles will be replaced by deterministic models.

useful, and it is an adequate approximation in parts of some problems. For example, gravity bombs can be dropped repeatedly from planes onto test ranges, and their accuracy can be measured. The tests can control for type of plane, type of bomb, altitude, speed, and the like. If these conditions are fixed, the bomb's error (called "ballistic dispersion") will still vary and the notion of statistical uncertainty will capture the variation adequately. If the same conditions can be fixed for future uses of the same type of bomb, and if they are *all* of the conditions that affect bomb error, then the notion of statistical uncertainty will be adequate to describe the predictive uncertainty about those future uses of the bomb. If the conditions cannot be fixed or some other condition can affect bomb error, then statistical uncertainty may not be adequate.

For example, suppose a bomb's deviation from its target is described by two components, one each for the errors in the northerly and easterly directions. It is common to use a bivariate normal (Gaussian) probability distribution to represent the variation in these two errors. Bivariate normal distributions are specified by five parameters: the means and variances of the two components, and their correlation. Statistical uncertainty about the bomb's accuracy can be broken into two parts: (1) uncertainty *about* the probability distribution from which the ballistic errors are drawn (e.g. uncertainty about the parameters of the error distribution), and (2) uncertainty *given* the probability distribution of the ballistic errors. The first of these uncertainties can be reduced by taking a larger sample in the test (that is, parameters can be estimated with greater precision), but the second one cannot. The latter is sometimes represented in military models and, less frequently, in operations research models generally, but the former is almost never represented.

State of the World Uncertainty[3]

Many uncertain things are either not repeatable, not observed or observable, or both. Hitch (1960) gives the example of the behavior of a sentient opponent in a game: In such instances as the escalation of a superpower conflict, the phenomenon is not repeatable and (fortunately) not observed. Whatever statements of uncertainty we may wish to make about such a situation, they cannot have the same

[3] Fisher (1971) used the term "state-of-the-world" uncertainty; but as with statistical uncertainty, others have used different names. Hitch (1960) used "genuine uncertainty" and Steinbrunner (1974) used "structural uncertainty."

basis as the statements we could make about bomb accuracy under specified conditions.

State-of-the-world uncertainty can creep into the problem of bomb accuracy when predictions must be extrapolated from conditions that have been tested to conditions that have not. Such extrapolations require a model relating the conditions to bomb error, and the accuracy of the model can be a central concern. In the Air Force spares and repair example, the models of part failures are used to extrapolate from peacetime flying programs to wartime flying programs that call for many more flying hours per aircraft. For this extrapolation, the state-of-the-world uncertainty can cover orders of magnitude.[4]

Behavior of people in the mechanism or the managed system adds a more complex layer to the state-of-the-world uncertainty. In World War II, aircrews were not generally reassigned to targets that they had missed and so lacked an incentive to follow the hazardous though effective bombing practices used in tests of bomb accuracy. Instead, the incentives were such as to incline them to safer but less effective techniques, which would diminish actual bomb accuracy.[5] For another example, field tests of Air Force units show that rates of part failures depend on the incentives acting on the pilots: If bombing accuracy is being measured, pilots have an incentive to report failures in their planes' fire control systems; but if accuracy is not measured, they have no such incentive, and reported failures drop.[6] Human interventions, even by people "on our side," introduce uncertainty that is often difficult or impossible to characterize.

The behavior of the people outside the managed system introduces still another layer of uncertainty. This state-of-the-world uncertainty has two variants. First, actions of those outside the managed system may be unpredictable. For example, before Pearl Harbor, many American military analysts believed that aerial torpedoes would be ineffective there, because they would probably hit bottom when dropped in such a shallow harbor. However, Japanese aviators developed a special fin that overcame the problem (Wohlstetter, 1962).

The interconnection of behavior within the system with behavior outside it may induce sequences of changes on both sides that are im-

---

[4]Donaldson and Sweetland (1968), Embry (1984), Pipp (1988), and unpublished RAND research by G.C. Crawford and M. Kamins and by M. Berman and T. Lippiatt all suggest that demands per flying hour could decrease by as much as a half in wartime for some important and expensive parts. These pieces of evidence are the subjects of a heated dispute, but it is clear that nobody has a grasp of what would happen.

[5]Personal communication from J. Stockfisch.

[6]Personal communication from H. Shulman.

possible to predict. For example, when the British installed centimeter radar on their anti-submarine planes, German submariners reacted by changing their tactics and diving when they sighted a plane. The British eventually deduced the depth the U-boats would reach by the time the plane reached them and began to set their depth charges accordingly, with devastating results. The Germans responded with the snorkel, which reduced the radar signature of their submarines. Even after the centimeter radar had been deployed, it would have been impossible to predict the outcome of this cycle of action and reaction.[7]

One implication of these examples is that analysts may not even be able to put bounds on plausible outcomes. Such bounds are essential to sensitivity analyses as commonly used; and without such bounds, the logic of sensitivity analysis must change. We discuss this further in Sec. V.

Such tools as subjective probability have been used to represent this kind of uncertainty, because subjective probability is manipulated according to the same rules as statistical uncertainty.[8] (Many proponents of subjective probability argue that statistical uncertainty is a special case of subjective probability.) Sometimes, state-of-the-world uncertainty extends to uncertainty about what the possible outcomes are, as in the centimeter radar case; and in these cases subjective probability may not be useful.[9] However, even leaving this objection aside, many who are not hostile to the idea of subjective probability do not accept its use for describing state-of-the-world uncertainty. They argue that a subjective probability distribution is

---

[7]Emery and Trist (1965) discussed four levels of complexity of interconnection between organizations and their environments. This paragraph is close to the environment they call a "turbulent field."

[8]Hodges (1987) surveyed the statistical literature on methods for handling uncertainty about model extrapolations. Methods involving subjective probability to greater or lesser extents were found to be the only formal methods that had been considered. Draper et al. (1987), using different terminology, surveyed the literature of formal methods for handling all the varieties of state-of-the-world uncertainty and made a similar finding. Alternatives to probability (e.g., fuzzy set theory, belief functions) were found to offer no solution to problems of extrapolation or specification and, thus, no advantage over subjective probability for this problem.

[9]Conrath (1967) distinguishes four degrees of uncertain situations: (1) a future outcome of interest is known; (2) the possible outcomes are known and a probability distribution across them is known, or, in subjectivist terms, is agreed upon; (3) the possible outcomes are known but a probability distribution across them is not known or agreed upon; and (4) the possible outcomes are not all known. The former three are states of "bounded uncertainty," and the last is a state of "unbounded uncertainty." Conrath applies these four degrees of uncertainty to all three parts of the classical decision problem: state of nature, action, and value (utility, loss).

an assertion of information like any other assertion; and if it is not founded on observations, on the basis of which other people might be persuaded to accept it, then it is useless and a distraction. For the most part, we adopt this view here.

EXAMPLES AND CRITIQUE OF THE TREATMENT OF UNCERTAINTY

In our discussion of policy analysis, we described three roles of analysis: operations analysis, systems analysis, and requirements analysis. The importance of statistical and state-of-the-world uncertainty varies depending on the role of analysis. In operations analysis, the system is operating and can be measured, so much of relevant uncertainty will be statistical. In systems analysis, the alternatives are largely hypothetical, so measurements are not available. In addition, many important aspects of the conditions and responsibilities will be hypothetical. Most, perhaps all, of the relevant uncertainty will be state-of-the-world. Requirements analysis falls between the other two roles. In the Air Force spares and repair example, the system is operational, but it must be outfitted against the possibility of a war. Thus, both kinds of uncertainty come into play.

Example 1: Air Force Spares and Repair Requirements

The responsibilities in the tasking are derived from guidance published by the Department of Defense and the services each year. This guidance contains information about several possible contingencies for each year of a five-year plan and directs which contingency is to be used to compute requirements. The Air Force produces a series of war mobilization planning documents based on the guidance, containing flying hours, sorties, force structures, projected attrition rates, deployment schedules, and other information for each squadron in the chosen contingency. This information is combined with detailed aircraft configuration data to estimate the tasking (flying hours per unit of time) that each aircraft part will experience. A similar tasking is derived for peacetime requirements, based on the forces' training needs.

In most Air Force requirement models, removals of parts are modeled as Poisson random variables with means proportional to flying hours, and the demand rate that specifies this proportion is derived from peacetime failure data. In some requirement models, the

Poisson model is replaced by a negative binomial model for which the variance-to-mean ratio is set in several ways.

Repair and transportation times are the most important elements in requirement models of the managed system. Usually these times are assumed to be unvarying or are exponential random variables. Management is assumed away by postulating ample (actually infinite) repair and transportation capacity.

Critique of Example 1

Requirements for spares and repair are computed using one approved contingency. This means that war mobilization planning documents do not address any of the state-of-the-world uncertainty associated with how the enemy might prosecute the war, with how U.S. commanders might design flying programs in response, with the effects of battle damage to bases, or with difficulties in operating the support system. Instead, resources are procured for essentially one possible war, in a fairly benign environment.

To illustrate the omission of these uncertain features of the problem, the effect of omitting battle damage from evaluations of parts stocks and alternative repair and resupply systems is being considered in ongoing but unpublished RAND research. An analysis of three airbases subjected each base to a typical enemy attack on the first day of the "war," and varying amounts of spare parts and repair capability were lost. The expected number of aircraft that would not be fully mission capable thirty days later was roughly twice the number projected in a tasking that omitted base damage but was otherwise the same. Plainly base damage is important and its extent is uncertain. If included it will affect the requirement, but it is ignored.

One obvious alternative to the single tasking is a group of taskings. Some readers have argued that a central purpose of the single war used in a requirement computation is to provide a "level playing field" for the various budgetary contenders, and to ensure that the computation produces a single number, not a range of numbers. But a specified group of taskings can provide as level a playing field as a single tasking and need not force the single requirement number to be replaced by a range of numbers. Instead, the single number could be tested against the range of taskings to see if it would suffice for all.

Some uncertainty does appear in the model of the spares and repair mechanism. It is all treated as if it were statistical uncertainty, captured in the Poisson or negative binomial model for part failures. This model posits several specific things, including constant average rate of demands per flying hour or sortie, with the same rate prevailing in peacetime and wartime. Even taking this model as accurate for the moment, the standard approach ignores one important aspect of statistical uncertainty—the uncertainty about the demand rates inserted in the model. However, the model is not accurate for many parts,[10] so its use for extrapolation suppresses uncertainty about predictions of failures wartime. In sum, the mechanism omits large and important state-of-the-world uncertainty, and its representation of statistical uncertainty has little to recommend it.

The same is true for the managed system. Both constraints and management are eliminated from the model of repair and management, when in fact both exist. The Department of Defense and the services have formal systems to ensure that priority is given to those needs that most affect the forces' ability to accomplish their missions. Although studies of priority rules show that their effect can be dramatic, these managerial adaptations and others that are known to exist are ignored in the requirements process.

Example 2: Army Munitions Requirements[11]

As with the Air Force spares and repair computation, the Army's munitions requirement calculation uses a single possible war. In addition, only one of the suite of models used has stochastic elements and is run repeatedly to generate variation in ammunition expenditure rates. Also, the munitions requirement uses the average of these rates and ignores variation. Moreover, data used as inputs to these models, notably kill probabilities, are themselves uncertain, sometimes very much so, but this is ignored. Finally, the models of warfare themselves are largely conjectural, because empirical study to build scientifically valid models has not and probably can not be done. The Army's analysts attempt to account for these difficulties by using conservative values of input data, but it is difficult to see how conservatism can be assessed in the face of such massive uncertainty.

---

[10]Thirty-five years of data consistently contradict the Poisson assumption for high-failure items. Mean demands are not constant even in peacetime, are not proportional to flying hours, and probably cannot be projected to wartime by a linear extrapolation. Evidence against the Poisson assumption for high-demand items is contained in Brown and Geisler (1954); Youngs, Geisler, and Brown (1955); Brown (1956); Astrachan, Brown, and Houghton (1961); Astrachan and Sherbrooke (1964); Hodges (1985); and Crawford (1988), among others. Evidence against constant mean demand rates and proportionality to flying hours is in Astrachan, Brown, and Houghton (1961), Hodges (1985), and Crawford (1988).

[11]See Girardini (1989).

Example 3: Air Force Munitions Requirements[12]

As with the Air Force's spares and repair requirement, a single contingency is used, along with official rates of aircraft attrition and similar model inputs. In contrast to the Army's calculation, some but not all determinations of munitions effectiveness account for statistical uncertainty. The discussion of uncertainty concludes with the judgment that adding more of these sources of statistical uncertainty "would add little or nothing to the plausibility of the results of a model operating within this framework"[13] (i.e., when so much state-of-the-world uncertainty is ignored).

Nonmilitary Examples

We cite some nonmilitary examples here to illustrate the wide-ranging nature of the deficiency in the treatment of uncertainty. Policy about low-level radiation hazard is made by administering high doses of radiation to lab animals, counting the resulting cancers, and extrapolating to low levels of radiation and to humans. Both parts of this extrapolation are highly uncertain, and that uncertainty is routinely ignored.[14] Quantitative risk assessment generally is plagued by such problems.[15] Fiscal policy is made using macroeconomic models, which are subject to great dispute and require projections of conditions that are conjectural at best.[16] Psychological and educational research make extensive use of so-called structural models, which rely on assumptions that are inherently unverifiable.[17] Even the determination of physical constants—as pristine as possible, one might think—is subject to uncertain nonexperimental error. International standards organizations are only beginning to regulate the reporting of such errors.[18]

---

[12] See Crawford (1989).

[13] Ibid., p. 42.

[14] Kalbfleisch and Prentice (1980); Hattis and Smith (1985); Freedman and Zeisel (1988).

[15] Hattis and Smith (1985); Hattis and Kennedy (1986); Kamins (1975); Speed (1985), among many others.

[16] Many examples are available; Greider's (1981) candid interviews with David Stockman are a good one.

[17] The Summer 1987 issue of the *Journal of Educational Statistics* (Vol. 12, No. 2) is devoted entirely to a critique of structural models by Freedman and to a discussion by defenders and other critics of such models. Holland (1988) shows the reliance on inherently unverifiable assumptions.

[18] Comité International de Poids et Mésures (1981).

IV. MECHANISM AND MANAGED SYSTEM ADAPT TO THE UNEXPECTED

If the common mode of analysis treats uncertainty so poorly, why aren't more military systems hopelessly bogged down? One reason is that managed systems and mechanisms a · not inert; they are systems that adapt to the unexpected. Instead of ignoring these facts or simply treating them as deficiencies in our models, we can use them as indications of how analysts can help managers cope with uncertainty by helping them adapt, instead of by helping them try to predict the unpredictable.

ADAPTABILITY OF THE MECHANISM AND MANAGED SYSTEM

The mechanism and managed system generally involve large human organizations. The people in them do not simply follow rules blindly, but respond in varying ways to events and to the incentives with which they are surrounded.[1] Although this fact may just make it more difficult to predict outcomes, if fully appreciated it may also reveal opportunities for designing adaptive systems.

To return to the Air Force example, some argue that the Air Force's model of part failures is deficient because it treats the mechanism as inert when it is not. Some types of equipment degrade instead of failing outright; because assessment of degradation requires human judgment, it can be highly subjective. Pilots and mechanics vary in the skill with which they diagnose faults. Because a pilot did not exercise some aircraft subsystem in his last mission, he may not observe that it has failed. Managers can change failure rates with decisions about what evidence should trigger a removal, and it is not unusual for a single failure to prompt wholesale inspections and removals.

This, too, is not just a problem for military analysts. For example, the Rational Expectations school of macroeconomic thinking has cast doubt on decades of models and derived policy with the hypothesis that agents in the economy anticipate the effects of government

[1]Hoenack (1983) gives a rich, if somewhat abstract, microeconomic construction of incentives created within organizations and the behavior they can be expected to induce.

macroeconomic policies and negate those effects by reacting to their expectations.[2]

Air Force spares and repair managers respond to the unexpected with a variety of adaptations. For example, they take parts from war reserve kits, remove parts from disabled planes to use in other planes, and schedule overtime work, to name a few. Similarly, Navy ships engage in "local barter," trading movies for coffee, for example. More formally, the services constantly reallocate material across units to maintain readiness: In December 1988 alone, the United States Air Forces in Europe (USAFE) responded to over 1500 redistribution requests for material needed by other USAF units worldwide.[3]

These adaptations are clearly useful, but they are rarely modeled and even more rarely evaluated in terms of their costs and contributions to overall system effectiveness. Instead, the Air Force's models postulate a system where management is unnecessary and has no value.

Again, this is not simply a phenomenon in military analyses. In a study of the New York City Fire Department, analysts found that when the rate of alarms got high, dispatchers adapted by dispatching fewer fire engines than their rules required, thereby keeping some engines in reserve.[4] The dispatchers took the chance that some fires would not initially receive enough units, but they retained some capacity to respond to new fires and to unexpected severity of existing fires.

GIVING MANAGERS MORE WAYS TO ADAPT

Rather than be chagrined that we do not model managers well, perhaps we should note that they are doing what prominent writers on systems analysis recommend, and help them do more of it. For example, Hitch (1960) noted that state-of-the-world uncertainties are omnipresent and usually irreducible and argued that the appropriate analytical tack is a shift "from a search for a better decision rule to a search for a better system [, f]rom sophistication in *judgment* to ingenuity in *design* [p. 5]. . . . Most of our relations are so unpredictable that we do well to get the right *sign* and order of magnitude of first differentials. In most of our attempted optimizations we are kidding our customers or ourselves or both. . . . It is much easier to find a sys-

[2]See, e.g., Begg (1982).

[3]Letter, Brigadier General Philip Metzler, USAFE DCS/Logistics, to Mr. John Abell, February 1989.

[4]RAND Fire Project (1979), pp. 431-432, p. 625, and sec. 11.3.3.

tem or strategy that dominates or nearly dominates some other system—say the one currently planned—than it is to find a system that dominates *all* other systems. . . . And if we can't *find* such a system, we can frequently invent one. In fact, it has been our experience that ingenuity is frequently more profitably exercised in invention than in mere judging" (p. 9).[5]

Analysts involved in selecting and evaluating the first geologic high-level radioactive waste repository, for example, spend tremendous effort to compute a probability that the proposed repository will not fail for tens of thousands of years.[6] These brave people attempt the impossible because their legislative and executive masters ask them to. The more pertinent problem is how the managers of the repository will respond to breaches, or how the design or site selection might be modified to make it easier to detect a failure and respond to it or, indeed, whether the time spent computing an uncomputable probability would be better spent on designing management responses.

Some readers have argued that management is only pertinent in operations analysis, not requirements or systems analysis. We do not agree. For example, the ability of managers to adapt is (or should be) at the heart of systems analyses of alternative high-level radioactive waste depositories. Similarly, the efficacy of management is the central issue in our discussion of the Air Force's spares and repair requirement. The next section elaborates this point.

THE EFFICACY OF MANAGEMENT: THE CLOUT STUDY

A study of a collection of Air Force spare parts initiatives called CLOUT[7] included both operational analysis—examining how repairs should be scheduled—and requirements analysis—computing quarterly and annual requirements for spares and repair. It added to the standard analysis both more uncertainty and more management

[5]In a similar vein, Ackoff (1964) presented six case studies of industrial operations research problems and drew the generalization that "greater improvement of performance than is obtainable by optimization may be obtained by changing the problem . . . what this can almost always do is make the system more *adaptive*; that is, make it more sensitive and responsive to changing conditions which affect its performance [p. 6]. . . . My hope is that we do not allow well formulated models and methods of solving them to give us a mental set which prevents our exploring more fundamental changes in the system than these models consider" (p. 11).

[6]For example, Eisenberg (1988), Doctor (1988), Booker (1988).

[7]CLOUT is an acronym for Coupling Logistics to Operations to meet Uncertainty and the Threat.

adaptations. Two examples of the additional uncertainty are a negative binomial model for part failures (instead of Poisson) and uncertain airbase damage arising from enemy attacks. The management adaptations included faster depot repair for critical parts ("priority repair") and allowing air bases to use the part stocks and repair capacity of nearby airbases ("lateral resupply and repair"), with all these choices made to maximize the number of aircraft available to fly missions. Both of these practices occur now to varying extents.

The effect of adding uncertainty to the standard analysis *without adding the management adaptations* was to double the expected number of aircraft that could not execute their missions one month into the war. When the management adaptations were added, the number of downed aircraft was reduced to slightly more than that obtained in the standard analysis.

For operations, one implication is that responsive scheduling of repairs can have a dramatic effect on aircraft availability in uncertain environments. A system to schedule depot repairs responsively has been developed and a prototype is now operating at the Ogden Air Logistics Center. Experience is showing that the gains projected in the CLOUT analyses are real.[8]

The implication for requirements calculation is that the computation is strongly affected by the inclusion of management. If the standard computation were changed by adding the sources of uncertainty considered in CLOUT, then the computed aircraft availability would be diminished. Availability could be brought back to more desirable levels by buying more spare parts, by managing the inventory of parts differently, or by some combination of these two. Each choice has costs and benefits that need to be evaluated as part of the requirement computation.

The problems of design and evaluation in the CLOUT study are examples of considerations generic to analyses that take realistic account of uncertainty. The next section describes in some generality these implications about the way analysts do business.

---

[8]Personal communication, John B. Abell.

V. IMPLICATIONS FOR ANALYSTS: DESIGN AND EVALUATION OF SYSTEMS

As mentioned in Sec. II, the five stages of a policy analysis are establishing the context, laying out the alternatives, predicting the consequences, valuing the outcomes, and making a choice. In the terms we have discussed, stage 2 is about system design, stages 3 and 4 are about system evaluation, and stage 1 is essentially setting the tasking. A proper recognition of uncertainty will affect all of these.

IMPLICATIONS FOR SYSTEM DESIGN: GENERIC STRATEGIES FOR UNCERTAIN SITUATIONS[1]

One message of the last section is that management is a resource with costs and benefits and is purchased in competition with other resources. This subsection illustrates some possibilities for system design and emphasizes the tradeoff between management and other resources by describing a collection of generic strategies for uncertain situations.

Passive and Active Strategies

Strategies for uncertain situations can be broken into passive and active groups. Passive strategies rely on size, scale, or diversity: Passive systems are *ample*; they *absorb*. Active strategies work on the premise that the future will consist of outcomes following from specific key events; the system must recognize those key events and *adapt* in a timely fashion with some activity relevant to the impending outcomes. Obviously, active strategies need resources that are ample in some sense, but they differ from passive strategies by adapting to do more with less. In this sense, passive strategies emphasize *goods*, and active strategies emphasize *management* of goods.

Passive strategies tend to be advantageous in stable situations where the uncertainty that is present can be well understood if not diminished, while active strategies tend to be advantageous in situations of great uncertainty where the likely error of predictions is either large or is itself quite uncertain.

[1]Madansky (1968) discussed some of the strategies listed here.

This discussion will now be elaborated with examples of three passive strategies and four active strategies.

Examples of Passive Strategies

Buy It Out. This strategy has many variants. First, one can cover all the possibilities. For example, the Maytag repairman doesn't know what tools he will need at a given house, but he doesn't need to because he carries them all in his truck. Second, it is possible to squeeze more out of what is already available. Scheduling overtime in repair shops is one example of this. Third, it may be possible to bank resources: Squirrels save nuts for winter, and people save money for retirement. In all these examples, the strategy is to have enough to cover whatever happens.

Operate on a Larger Scale. An insurance company guards against losses from a single person's claims by grouping that person with many similar people. Because the group is less erratic than the individuals, no individual needs to pay enough money to cover the largest insured claim. Auto repair chains in large cities use a similar strategy when they have a central warehouse for most of their parts and dispatch parts from the warehouse as needed, instead of trying to buy enough parts to stock each repair shop against all eventualities.

Diversify. An insurance company holding many fire insurance policies in one area is subject to sudden bankruptcy if all the policyholders lose their homes to a large fire. The company guards against this by reinsurance. Investors do the same thing when they diversify their portfolios.

Examples of Active Strategies

Get More Information. One variant on this strategy is to procrastinate and react as late as possible. Procrastination preserves options, although it can incur costs. Quarterbacks use this strategy when they call audible plays at the line of scrimmage. A similar strategy is practiced in the Services' depot repair organizations: While they develop quarterly repair plans to meet the force's projected needs, they adjust those plans somewhat as the force's actual needs become known. For example, if demands for a part are so great that wartime capability is threatened, the depots rush unplanned repairs through the shops. Another variant of this strategy is to buy more information without necessarily procrastinating first. Providing current information about part inventories to a responsive depot repair system is an example. Another was suggested in the example of the repository for high-level radioactive waste: Aggressively monitor the performance of the containment structure, and respond to it.

Use General or Flexible Resources. One particularly effective way to exploit information relies on obtaining and maintaining flexible general-purpose resources. For example, the effectiveness of the quarterback's audible call depends on the ability of the other team members to memorize and execute a large number of plays. The CLOUT options, by contrast, are of little use to a repair shop that can repair only one item. In munitions procurement, flexible manufacturing provides a capacity to produce several different munitions depending on what the unfolding conflict demands. Aircraft leasing firms have a general resource in their planes. Although an airline trying to lease a single superfluous plane has just one plane, and information is useful to it only to the extent that it helps the airline lease that single plane, a leasing firm has several planes, and its information is relevant to its whole fleet. Flexibility can also be obtained by selection. In World War II, the U.S. Army trained new troops into the branches based on the requests from the fielded units. It did not simply train recruits according to preset schedules or force structures, it adjusted training as the units began to understand their needs.

Push the Risk onto Someone Else. Airlines face an uncertain world under deregulation and push some of the resulting risk onto leasing firms. The leasing firms own the planes and worry about who will use them, not the airlines. Just-in-time inventory systems push the risk of supply onto the supplier. Wholesalers and job shops accept such risks by adopting one or more of the strategies described earlier. In general, the person or organization accepting such risk must have some advantage if this strategy is to work. The leasing firm's advantages are its flexible resources—aircraft that can be used anywhere—and its market information. The wholesaler's advantages are its scale and its knowledge of recent or pending demands. The job shop's advantages are its flexible resources.

Combine Strategies. Munitions manufacturing for wartime combines at least three strategies. The first is surge capability, the ability to push more out of existing production lines with overtime and multiple shifts. This is a "buy it out" strategy. The second is flexible manufacturing. The third is mobilization of mothballed capacity, a diversification strategy. The choice of which form of manufacturing, when, in what quantities, and for what munitions would be made during a mobilization and during the war itself, depending on what munitions were being consumed, which were effective in combat, which were cost effective, and so on.

The CLOUT initiatives for Air Force spares and repair are another example of a combined strategy: They involve lateral repair and resupply within a theater of operations (greater scale), current information about inventories and planes awaiting parts (buying more information), and adaptively scheduling repair and distribution to maximize aircraft availability (procrastinate and react).

Passive and Active Strategies Compared

The foregoing discussion was not an attempt to prescribe precise solutions for any problem, but a collection of analogies to make a few points. One central feature of strategies for uncertain situations is clear: The cost and effectiveness of the overhead needed for an active system must be balanced against the cost and effectiveness of buying amply for a passive system.

On the cost side, the main burden of a passive strategy is buying enough "stuff," whatever that happens to be in the case at hand. Overhead is fairly simple and thus cheap. Management is clearly a more important resource in diversification strategies than in "buy it out" strategies, but only for the initial decisions on whether and how to diversify; but in active strategies, management is the most important resource. It is needed to recognize key events and adapt to them, and it costs: Overhead is more elaborate and expensive, particularly for information systems.

On the effectiveness side, the ampleness of a passive strategy is wonderful when it can be had, but it must be possible to define ampleness, to attain it, and to afford it. For example, when airbase attacks can destroy repair capacity and stocks of parts, no amount of extra parts at the base will suffice to cover this possibility. In this case, ampleness cannot even be defined. In other cases, ampleness can be defined but it cannot be attained. It would be desirable, for example, if every aircraft mechanic could perform every repair task, but a mechanic can learn only so many specialties. Even when ampleness can be defined and attained in principle, it may be too expensive: Uncertainty may be so great that excessive quantities must be bought to reach an acceptable level of confidence.

If ampleness cannot be achieved for any of these reasons, active strategies may be more effective. An active system must be able to recognize key events, compute an advantageous reallocation of resources, and execute it. If any of these capabilities is impaired, or if there are no key events to recognize or no advantageous reallocations, the effectiveness of an adaptive strategy is impaired.

These generic strategies may be useful as a source of ideas for analysts designing possible systems. Even more useful would be finding out what actual managers do to adapt in real situations. History offers an abundance of real managers in real wars.[2] In addition, exercises are run now in the armed forces; these could be studied, or exercises could be run specifically to stress organizations and observe how managers adapt.[3] Finally, analysts can go watch managers in their day-to-day activities. The adaptive repair and distribution policies analyzed in the CLOUT study would extend and formalize similar things that Air Force managers do now, in peacetime.

IMPLICATIONS FOR SYSTEM EVALUATION: SENSITIVITY ANALYSIS IS NOT ENOUGH

Like any other system, a responsive system must be evaluated and compared with alternatives, and an acknowledgment of uncertainty and adaptation has implications for evaluation. Here, we argue that more use of such standard tools as probability and sensitivity analysis will be helpful in some situations, but those situations occur infrequently. In other situations, a different kind of logic and analysis is needed to connect the arithmetic in the analysis to the actual problem.

What Standard Tools Can Do

Analysts may be able to use familiar tools more often to take better account of uncertainty. It may be possible to accurately represent statistical uncertainty in models by replacing some constants with probability distributions and to represent some management adaptations in models. For instance, in the Air Force spares example, it is possible to replace the Poisson failure model with a negative binomial model that has variability more like that observed in peacetime and to represent cannibalization of parts from disabled aircraft. It may also be possible to use more sensitivity analysis for uncertain things about which probability distributions are difficult to specify or justify.

[2]For example, see "War of the Accountants," in Van Creveld (1977).

[3]Pipp (1988), for example, describes the Air Force's Coronet Warrior exercise, in which the Dyna-METRIC model was tested. Among other things, logistics personnel were observed and found to be more innovative and to perform better than predicted.

In the spares example, different levels of airbase damage can be examined in a sensitivity analysis (and were in the CLOUT analysis).

These standard textbook devices will be useful and adequate when the model of statistical uncertainty and management adaptations have a firm grounding in fact and when the things to be varied in the sensitivity analysis can be bounded with certainty. In such cases, plenty of tools exist and the logic connecting the analysis to the world is straightforward: *If* all the important uncertain things can be bounded with certainty, and *if* variation within those bounds does not affect the outcome of the analysis, *then* uncertainty about those things will not upset the results of the analysis.

What Standard Tools Can't Do

We believe the situation just described does not hold very often, and we are only selectively sanguine about the utility of spending a great deal of extra effort on adding probability distributions to models and doing dozens of sensitivity runs. Indeed, we doubt that it is ever possible to represent all the salient uncertainty, although it may be possible to represent enough of it. Even in operational analyses, in which data are available, models of statistical uncertainty may be difficult or impossible to justify. This is the case in spare parts work. Even for peacetime, part failures change from period to period in ways that are difficult to reconcile with any handy statistical model. Further, the things that are candidates for sensitivity analyses often cannot be bounded with anything resembling certainty. Extrapolations of part failures to wartime are simply unknown, and experience from earlier wars indicates that our predictions may be wrong by orders of magnitude.

One can still spend considerable effort building more statistical variation into models and doing dozens of variations on them in sensitivity analyses, and this is commonly done. But without the solid grounding for the model of statistical uncertainty and certain bounds on the things varied in the sensitivity analysis, the analysis is without an explicit logical connection to the world, and the extra effort is not of clear value.

So what else can be done when these textbook tools do not do the job? What can fill the space between the model/analysis and the world about which it is to be informative? The space can be filled in logically clean and unclean ways. The term "logically unclean" is not meant to convey condemnation; it is meant to convey that although a logically unclean argument may be a necessary evil, it is still undesirable. Logically unclean connections between a model/analysis and the world are necessarily more subtle and elaborate than logically clean connections; we do not understand them well enough to pursue them here and will defer consideration to another study.

Two Logically Clean Alternatives[4]

Two logically clean connections are readily available and often useful. The first is to implement a proposed system, perhaps as a prototype, and see if it performs adequately. As noted above, part of the CLOUT system is being prototyped, and it appears to perform roughly as advertised *for the peacetime conditions in which it is being tested*. Given that it does perform adequately, it is almost irrelevant if the models and assumptions inside the system are accurate: If the system works, it works, and it is not necessarily cost effective to replace inaccurate models with more accurate ones.

Tests like this, however, are available for only some problems, generally for operational analyses involving adjustments to existing systems. Other problems may be susceptible to the other logically clean argument, an a fortiori argument. For our purpose, a fortiori arguments can be described as follows: If condition X were true, then policy A would be preferable to the other candidates, but the actual situation deviates from X in specific ways that favor policy A even more, so a fortiori policy A is preferable. The CLOUT analyses are animated by such an argument. The model used in these analyses treats part failures as more predictable than they really are and allows fewer management adaptations than are likely to be available. In this world, the CLOUT initiatives pay. In an actual war, conditions will probably be more favorable to CLOUT, so a fortiori CLOUT appears to prevail over the current arrangement.

The argument has three parts—that X implies A is preferable, that X represents a boundary on the actual situation, and that deviations from X will favor A. Each part of the argument is essential and may force certain things on analysts using it. It may be possible to assert that "X implies A is preferable" only by means of a model. The introduction of the model may, because of its unrealism, undermine the assertion, but it need not. For example, the Army's JANUS combat simulation model has been attacked because of many patently unreal-

[4]Neither of these alternatives is our invention, or even new. We emphasize them to focus attention on the logic of models and analysis, especially in the face of pervasive uncertainty. Our emphasis on the logic of the argument is unusual if not novel.

istic features of its simulation of the behavior of individual weapons, sensors, and their operators. But JANUS can still be used to drive an a fortiori argument. Specifically, JANUS is a benign environment for notional systems: They are operated properly and work as advertised, they are generally not subjected to countermeasures designed specifically for them, and they are not subject to Murphy's Law. Thus, if proposed systems don't pay in JANUS, a fortiori they will not pay in real combat.[5]

It may not be possible to find a set of bounding conditions X that permits an a fortiori argument. For example, many have argued that models simulating the European theater of war are so complex that a set of bounding conditions is too much to ask for. Indeed, some argue[6] that the Defense Department's reliance on a fortiori arguments using worst-case taskings for Europe are so extreme that they conceal useful and feasible gains. Further, it may be a matter of contention how actuality differs from X and whether the deviations favor A. Some have attacked the CLOUT analysis with the assertion that wartime will not necessarily be more favorable to CLOUT, because of degradations in the necessary command, control, and information systems. This is a real concern and the subject of current study. Finally, if an acceptable model relating X to outcomes is available, it may be necessary to use it to understand whether possible deviations from X will indeed favor A.[7]

IMPLICATIONS FOR SETTING THE TASKING

A tasking can serve several purposes. It may represent an actual prediction of conditions, or it may serve only to constrain the analyses of claimants in a budget exercise. In either role, uncertainty can be salient and can affect the form the tasking takes. If the tasking represents an attempt to predict, then generally it should represent uncertainty. It may do this with multiple sets of responsibilities or conditions, whose use will depend on the nature of the analysis they feed. If the tasking is a constraint on the analyses of budget claimants, it can still consist of several sets of responsibilities or conditions, with the injunction that the budget claims should perform adequately against each part of the tasking.

[5]This observation is due to Richard Salter.

[6]See, *for example*, Davis (1988).

[7]In general it will be easier to construct a fortiori arguments in operational analyses than in other kinds of analysis, because fewer features of the problem are subject to change, and the behavior of the operational system should be more clearly understood.

We have few specifics to offer now about selecting the multiple responsibilities and conditions for a tasking. We conjecture that progress will come from the idea that the tasking should be set up to drive a fortiori arguments. It may be possible to use a tasking to create incentives for the tasked organization to institutionalize management adaptations. For example, the Air Force Logistics Command might task its subordinate organizations with (say) five responsibilities/conditions, with the objective of meeting a minimum level of aircraft availability in each, and with a kicker: If the subordinate organization can implement an adaptive procedure and demonstrate a saving in current operations, the organization keeps a percentage of the saving. We mention this example not because it is necessarily practicable, but only to suggest the possibilities of further work on setting taskings

REFERENCES

Ackoff, R. L., "Beyond Optimization," *Opsearch*, Vol. 1, No. 1, 1964, pp. 5–11.

Astrachan, M., B. B. Brown, and J. W. Houghton, *A Comparative Study of Prediction Techniques*, The RAND Corporation, RM-2811-PR, September 1961.

Astrachan, M., and C. C. Sherbrooke, *An Empirical Test of Exponential Smoothing*, The RAND Corporation, RM-3938-PR, March 1964.

Begg, D.K.H., *The Rational Expectations Revolution in Macroeconomics*, Johns Hopkins University Press, Baltimore, 1982.

Booker, J. M., "Combining Qualitative and Quantitative Data: The Use of Subjective Judgment Modeling," *Section on Physical and Engineering Sciences, Proceedings of the 148th Annual Meeting of the American Statistical Association,* Washington, D.C., 1988.

Brown, B. B., *Characteristics of Demand for Aircraft Spare Parts*, The RAND Corporation, R-292, July 1956.

Brown, B. B., and M. A. Geisler, *Analysis of the Demand Patterns for B-47 Airframe Parts at Air Base Level*, The RAND Corporation, RM-1297, July 1954.

Builder, C. H., *The Masks of War: American Military Styles in Strategy and Analysis*, Johns Hopkins University Press, Baltimore, 1989.

Comité International de Poids et Mésures, *Proces-Verbaux des Séances*, Vol. 49, 70^e Session, 1981, p. A1.

Conrath, D. W., "Organizational decision making behavior under varying conditions of uncertainty," *Management Science*, Vol. 13, No. 8, April 1987, pp. B487-B500.

Crawford, G., *Variability in the Demands for Aircraft Spare Parts*, The RAND Corporation, R-3318-AF, January 1988.

Crawford, G., *The Air Force's Munitions Requirements Process (The Nonnuclear Consumables Annual Analysis)*, The RAND Corporation, N-2821-P&L, March 1989.

Davis, P. K., *The Role of Uncertainty in Assessing the NATO/Pact Central Region Balance*, The RAND Corporation, N-2839-RC, December 1988.

DeGroot, M. H., "A Conversation with Persi Diaconis," *Statistical Science*, Vol. 1, No. 3, 1986, pp. 319-334.

Doctor, P. G., "Statistical Issues in the Validation of Performance Assessment Modeling," *Section on Physical and Engineering Sciences, Proceedings of the 148th Annual Meeting of the American Statistical Association*, Washington, D.C., 1988.

Donaldson, T. S., and A. F. Sweetland, *The Relationship of Flight-Line Maintenance Man-Hours to Aircraft Flying Hours*, The RAND Corporation, RM-5701-PR, August 1968.

Draper, D., J. S. Hodges, E. E. Leamer, C. N. Morris, and D. B. Rubin, *A Research Agenda for Assessment and Propagation of Model Uncertainty*, The RAND Corporation, N-2683-RC, November 1987.

Eisenberg, N. A., "The Department of Energy's Perspective on Performance Assessment Modeling," *Section on Physical and Engineering Sciences, Proceedings of the 148th Annual Meeting of the American Statistical Association*, Washington, D.C., 1988.

Embry, L. B., *Estimating Wartime Support Resource Requirements: Statistical and Related Issues*, The RAND Corporation, N-2163-AF, July 1984.

Emery, F. E., and E. L. Trist, "The causal texture of organizational environments," *Human Relations*, Vol. 18, 1965, pp. 21–32.

Fisher, G., *Cost Considerations in Systems Analysis*, Elsevier, New York, 1971.

Freedman, D. A., and H. Zeisel, "From mouse to man: The quantitative assessment of cancer risks," *Statistical Science*, Vol. 3, No. 1, February 1988, pp. 3–56.

Girardini, K., *The Army's Conventional Munitions Acquisition Process*, The RAND Corporation, N-2864-P&L, February 1989.

Greider, W., "The Education of David Stockman," *The Atlantic Monthly*, December 1981, pp. 27–54.

Hattis, D., and D. Kennedy, "Assessing risks from health hazards: An imperfect science," *Technology Review*, May/June 1986, pp. 60–71.

Hattis, D. and J. A. Smith, "What's wrong with quantitative risk assessment?" in J. M. Humber and R. F. Almeder (eds.), *Quantitative Risk Assessment* in *Biomedical Ethics Review 1986*, Humana Press, Clifton, New Jersey, 1987, pp. 57–105.

Hitch, C., "Uncertainties in Operations Research," The RAND Corporation, P-1959, April 1960.

Hodges, J. S., *Modeling the Demand for Spare Parts: Estimating the Variance-to-Mean Ratio and Other Issues*, The RAND Corporation, N-2086-AF, May 1985.

Hodges, J. S., "Uncertainty, Policy Analysis, and Statistics," *Statistical Science*, Vol. 2, No. 3, 1987, pp. 259-291.

Hoenack, S. A., *Economic Behavior Within Organizations*, Cambridge University Press, New York, 1983.

Holland, P. W., "Causal inference, path analysis, and recursive structural equations models," in C. C. Clogg (ed.), *Sociological Methodology 1988*, Vol. 18, American Sociological Association, Washington, D.C., 1988, pp. 449-484.

Kalbfleisch, J. D., and R. L. Prentice, *The Statistical Analysis of Failure Time Data*, Wiley, New York, 1980.

Kamins, M., "A reliability review of the Reactor Safety Study," The RAND Corporation, P-5413, April 1975.

Madansky, A., "Uncertainty," Chapter 8 in E. S. Quade and W. I. Boucher (eds.), *Systems Analysis and Policy Planning: Applications in Defense*, Elsevier, New York, 1968.

Pipp, Captain D. C., "Coronet Warrior: a WRSK flyout," *Air Force Journal of Logistics*, Vol. 12, No. 3, Summer 1988, pp. 1–4.

Quade, E. S., *Analysis for Public Decisions* (2d ed.), North-Holland, New York, 1979.

RAND Fire Project, *Fire Department Deployment Analysis*. North Holland, New York, 1979.

Speed, T. P., "Probabilistic risk assessment in the nuclear industry: Wash-1400 and beyond," in L.M. LeCam and R.A. Olshen (eds.), *Proceedings of the Berkeley Conference in Honor of Jerzy Neyman and Jack Keifer,* Wadsworth, Belmont, California, 1985, pp. 173–200.

Steinbrunner, J. D., *The Cybernetic Theory of Decision*, Princeton University Press, Princeton, 1974.

Stokey, E., and R. Zeckhauser, *A Primer for Policy Analysis*, W. W. Norton and Co., New York, 1978.

Wohlstetter, R., *Pearl Harbor Warning and Decision*, Stanford University Press, Stanford, California, 1962.

Van Creveld, M., *Supplying War*, Cambridge University Press, New York, 1977.

Youngs, J.W.T., M. A. Geisler, and B. B. Brown, *The Prediction of Demand for Aircraft Spare Parts Using the Method of Conditional Probabilities*, The RAND Corporation, RM-1413, January 1955.