AD 625 947

THE CHOICE OF ANALYTICAL TECHNIQUES IN COST-EFFECTIVENESS ANALYSIS

Alfred Blumstein

Institute for Defense Analyses
Arlington, Virginia

October 1965

Research paper P-206

REPORT

selection aids

Pinpointing R & D reports for industry

Clearinghouse, Springfield, Va. 22151

U.S. GOVERNMENT RESEARCH AND DEVELOPMENT REPORTS (USGRDR)---SEMI-MONTHLY JOURNAL ANNOUNCING R&D REPORTS. ANNUAL SUBSCRIPTION $30.00 ($37.50 FOREIGN MAILING). SINGLE COPY $3.00.

U.S. GOVERNMENT RESEARCH AND DEVELOPMENT REPORTS INDEX---SEMI-MONTHLY INDEX TO U.S. GOVERNMENT RESEARCH AND DEVELOPMENT REPORTS. ANNUAL SUBSCRIPTION $22.00 ($27.50 FOREIGN MAILING). SINGLE COPY $3.00.

FAST ANNOUNCEMENT SERVICE---SUMMARIES OF SELECTED R&D REPORTS COMPILED AND MAILED BY SUBJECT CATEGORIES. ANNUAL SUBSCRIPTION $5.00, TWO YEARS: $9.00, AND THREE YEARS: $12.00. WRITE FOR AN APPLICATION FORM.

DOCUMENT PRICES---ALMOST ALL OF THE DOCUMENTS IN THE CLEARINGHOUSE COLLECTION ARE PRICED AT $3.00 FOR PAPER COPIES AND 65 CENTS FOR COPIES IN MICROFICHE.

COUPONS---THE CLEARINGHOUSE PREPAID DOCUMENT COUPON SALES SYSTEM FOR PURCHASING PAPER COPIES AND MICROFICHE PROVIDES FASTER, MORE EFFICIENT SERVICE ON DOCUMENT REQUESTS. THE PREPAID COUPON IS A TABULATING CARD WITH A FACE VALUE OF THE PURCHASE PRICE OF A CLEARINGHOUSE DOCUMENT ($3.00 PAPER COPY OR 65 CENTS MICROFICHE). IT IS YOUR METHOD OF PAYMENT, ORDER FORM, SHIPPING LABEL, AND RECEIPT OF SALE.

COUPONS FOR PAPER COPY (HC) DOCUMENTS ARE AVAILABLE AT $3.00 EACH OR IN BOOKS OF 10 COUPONS FOR $30.00. COUPONS FOR MICROFICHE COPIES OF CLEARINGHOUSE DOCUMENTS ARE AVAILABLE IN BOOKS OF 50 COUPONS FOR $32.50. WRITE FOR A COUPON ORDER FORM.

U. S. GOVERNMENT PRINTING OFFICE : 1968 O - 328-773

RESEARCH PAPER P-206

THE CHOICE OF ANALYTICAL TECHNIQUES IN COST-EFFECTIVENESS ANALYSIS

Alfred Blumstein

October 1965

IDA INSTITUTE FOR DEFENSE ANALYSES
RESEARCH AND ENGINEERING SUPPORT DIVISION

Contract SD-50
In-House Research Study

CONTENTS

FIGURES

I. INTRODUCTION

This paper was presented at the Cost-Effectiveness Symposium of the Washington Operations Research Council held in Washington, D. C., June 14 to 16, 1965. Discussion in this paper is confined to an indication of the kinds of models that might be used to determine a system's effectiveness, how these models might be used, and some of the considerations that go into the selection of a particular model or models.

Other factors, such as the question of criterion, consideration of various kinds of measures of a system's effectiveness, and the problem of determining costs, were covered by other speakers at the Symposium.

II. KINDS OF MODELS

The problem to be discussed is the assignment of a measure of effectiveness to a system's performance. With that, we can then choose from among a class of alternatives the system with the maximum effectiveness for a given cost, or with a minimum cost for a given effectiveness.

Ideally, we would like to examine the performance of our alternative systems in the real world. But, for many obvious reasons, such as excessive expense, the unavailability of the system, excessive danger, etc., we cannot operate in the real world. Thus, we revert to a model world where we create a model or abstract representation of the real system with which we can manipulate and experiment. Then, to the extent that the model is an adequate representation of the real world for our purposes, conclusions we draw in the model world as a result of our manipulations may be extrapolated to the real world.

I would like to characterize these models along the spectrum of abstraction, as shown in Fig. 1. First, we start out with the real world system which we do not or cannot work with. Then, we examine models of it. The closest approximation to this would be an operational exercise using the systems in question. This might be the conventional military exercise in which troops cover the fields of North Carolina, or ships cover a major fraction of the Atlantic Ocean. But it's different from a real military operation in that the enemy is simulated, the danger of enemy action need not be considered and special precautions are introduced to avoid accidents. In an air traffic control situation, an operational exercise might involve using professional air traffic controllers in a real tower controlling flying airplanes; but, here again, safety rules are imposed to prevent a collision in case the system fails. The controller knows that these precautions exist, and his behavior is affected. The operational exercise, then, is a very close approximation to the real system but contains constraints that would not normally exist in the real world. The cost of running such an exercise is very large, and the number of cases which can possibly be examined is small.

The next region along this spectrum is gaming. Here, we remove from the representation of the real world those components that can most easily be simulated by a simple analog. For example, we represent the troops and aircraft by some means which retains their few most important attributes for the system under consideration. Most importantly, for systems with important human decision components, we retain the man to make the decision; that portion of the system is the most difficult to represent otherwise. In an air traffic control context, for instance, we would have a controller watching a real radar scope, as in the operational exercise, but the blips on the scope would now be generated by an electronic blip generator manned by "pilots" turning the appropriate knobs upon instruction from the controller. Since the controller's decisions are the crucial part of the representation, we try to make his information input and output as precise as possible, but we avoid the considerable expense of flying airplanes through the sky.

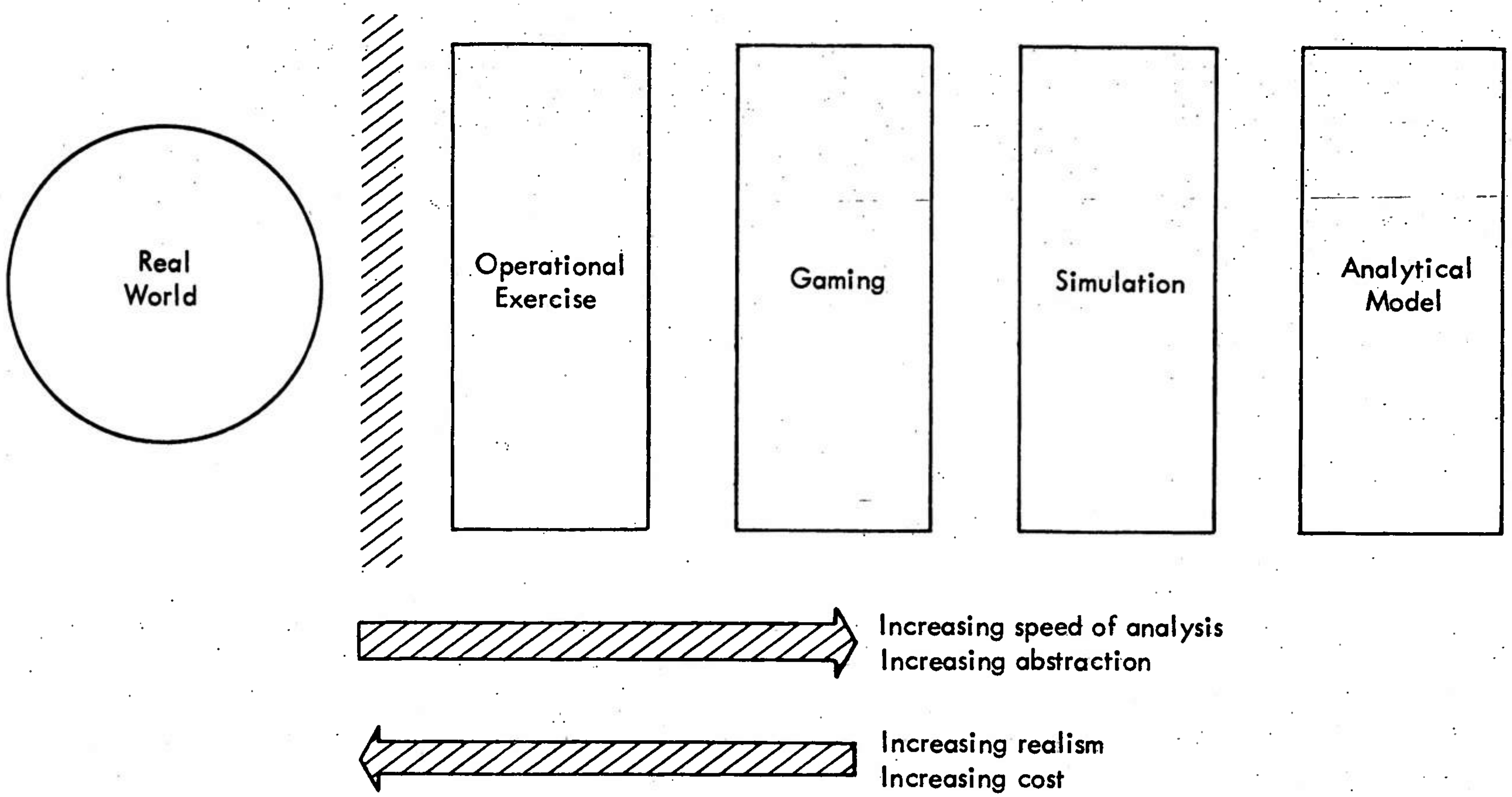

FIGURE 1 Spectrum of Models

The next stage of abstraction is to remove even the human decision-maker from the representation. At this point we can simulate the operation of our system on a digital computer in which the human decision rules are explicitly programmed--albeit with stochastic elements. Removing the human from the representation, then, permits operation at a rate considerably in excess of real time and data can be generated far faster than with humans making the decision. We pay the penalty, however, that our representation cannot reflect the flexibility and versatility inherent in a human decision-maker.

So far, in all the classes of models we have considered, there has been a one-to-one correspondence between time in the real world and time in the model. In the last class of models that we describe, analytical models, we abandon even this relationship. Here we generate a system of equations that relates characteristics of the system on one hand, to measures of effectiveness on the other. Examples of this are linear programming models, queueing models, inventory models, and the whole class of systems of equations. Here, in the air traffic control context, for instance, we might use a queueing model to determine the delay imposed on aircraft as a function of various priority rules, arrival rates of aircraft and service capabilities of alternative control systems. These models have all the virtues inherent in mathematics--the opportunity to manipulate the models and find maxima or minima, the opportunity to explore parametric relationships, and the consequent generalizability of the results.

So, in examining the spectrum, we see as we go from left to right that we have increasing abstraction from the real world or alternatively, increasing realism or closer approximation to the real world as we go to the left. On the other hand, the cost goes up as we get more realistic, and the rate at which we can examine different situations goes down. For reasons of economy and completeness of analysis, we would like to operate as far to the right on this spectrum as possible -- based on the limitations of our ability to represent the system adequately with that degree of abstraction. In general, each of these

Increasing complexity →

Model \ Problem	Strategic Lift	Strategic Offense and Defense	Tactical Air and Navy	Ground Combat	Counter-insurgency
Real World					X
Exercise				X	
Gaming			X	X	
Simulation		X	X		
Analysis	X	X			

Increasing abstraction ↓

FIGURE 2 Model–Problem Relationships

classes of models has its role in an analysis; in the best of all worlds, we would like to make use of all points on the spectrum. We can use the more abstract, faster systems to examine a large range of possibilities, identify the most important regions in that range, and then examine those points in more accurate detail with the more realistic models at the left. On the other hand, we use the more realistic model to help provide information for the analytical models at the right. We use them, for instance, to provide input data for the more abstract models on the right, and as a basis for formulating the models.

III. RELATIONSHIP OF MODELS TO DEFENSE SYSTEMS

I would now like to relate these classes of models to various classes of military systems. First, I would like to use William Niskanen's ranking of military problem areas in terms of their ease of analysis. (See columns of Fig. 2.) The easiest, most nearly deterministic military system is that designed to provide strategic lift. Behind that is the problem of spasm-type strategic offense and then defense, where there is considerably greater uncertainty, but where a scenario can be defined relatively well with little human decision of a military type entering into it. Beyond these areas in difficulty are those of tactical air warfare and naval engagements in which the field commanders continually make a wide variety of significant tactical decisions. Of even greater difficulty is the analysis of ground warfare where the range of decisions open to the field commanders is much larger, the results are much more sensitive to these decisions, and considerations of troop morale, will to fight, details of ground position, terrain, etc., have a much more significant effect on the outcome of an individual engagement. Far more difficult to analyze are the problems of counterinsurgency in which the military aspects of the conflict represent only a relatively small part of the battle. There is little known or understood about the considerations affecting the outcome of this type of conflict, and even the sign of the effect is very uncertain.

Aside from the increasing difficulty of determining the relationships in these classes of systems, they are further characterized by an increasing difficulty in the identification of even the appropriate non-cost measures of effectiveness. In strategic lift, the considerations are clearly load-carried and time-to-carry. In the strategic systems, we might use damage inflicted on an enemy or self-damage avoided, but even here we start to feel uneasy about whether the whole problem is reflected in such simple criteria. In tactical air warfare, we normally use a damage criterion, but our feeling of unease increases since target selection, timeliness and self-damage avoided in our mission all become relevant. In ground combat, the situation is further complicated by considerations of position, integrity of forces, logistic implications, and even the effects of the outcome of an engagement on the troops' morale. Finally, in counterinsurgency, since the war is fought over the well-known "minds and hearts" of the people, it becomes very difficult indeed to find proxy measures of effectiveness relating the performance of a military system to such considerations.

We can examine in Fig. 2 the relationship between these classes of military systems and the classes of models. We see that the more complex the military situation, the less we can accept the abstraction in the model. With strategic lift we can examine the whole problem analytically through a complex linear program; for instance, we can impose the delivery constraints and select aircraft and schedules to minimize the cost of performing that mission or minimize the time with resource constraints. In the strategic situation the stochastic aspects appear far more prominently. Since these aspects are far more difficult to treat analytically, we can treat only a relatively simple portion of the problem analytically and we normally find that we must move into a more complex simulation. The simulation allows us the flexibility of bringing in whatever form of statistical distributions we wish. It also allows us to bring in many of the interactions which must be assumed out in an analytical model. As I indicated before, however, we can let the simulation and the analytical formulation interact so that the simulation serves as a check against the analytical formulation and provides input information to it. Thus, the analytical formulation permits us to identify the critical pieces which must be examined in more complete detail on the simulation.

In the tactical air area we are almost forced into a simulation for any meaningful treatment of a problem, although we may sometimes have an analytical formulation of a very small piece of the problem. In many cases, particularly where significantly new weapons systems are involved, we want to determine engagement rules, firing doctrine, and so on, through experimentation with human subjects making the decisions. This brings us into a slower, more expensive realm of gaming because the commanders' decisions assume greater importance.

In ground combat, only the most simplified aspects of the problem can be treated analytically. We can try to simulate, but the complexity of the ground commanders' decisions often forces us into gaming where we use military officers to make the tactical decisions. Here, operational exercises become very important to provide the data inputs that will be used in the games.

In counterinsurgency gaming, we are neither very sure what kinds of decisions are meaningful nor certain about how to conduct an operational exercise. So, to a large extent, we are forced at this stage of our understanding to rely almost exclusively on what is going on in the real world to gain some understanding of the process. We can try to use the other modeling techniques for segments we can isolate, but models are applicable there more for structural understanding than for selection decisions.

IV. USES OF THE MODELS

In general, these models are very useful in that merely the attempt to formulate the model systematizes our thinking about the systems, raises questions of compatibility, and forces a structuring of the system in its operational context. In a cost-effectiveness context, in particular, the models can be used in two primary ways: for optimization and for comparison.

First, I would like to differentiate between two classes of system alternatives: those that differ in degree and those that differ in kind. Two systems that differ in degree have the same structural relationships among their components, but the particular parameter values differ, that is, they can be described by the same set of mathematical relationships, but the x's take on different values. Two systems that differ in kind, then, have different structural relationships. For example, an air traffic control system based on distance separation differs in kind from one based on time separation. Two distance-separation systems, one with a three-mile separation and another with a five-mile separation, differ in degree.

For systems which differ only in degree, and where there is a single measure of effectiveness (which may be a combination of several measures of effectiveness) the model can be used to select the optimum parameter values among the infinitude of such systems.

Among systems which differ in kind, one can use the models only for comparing the alternative systems, and then choose the best among the ones compared. But it must be recognized that the "optimum" or best possible cannot be chosen thereby, since there is always the possibility that another system differing in kind might exist which is better than those compared.

V. CONSIDERATIONS IN MODEL SELECTION

At this point, some of the considerations that go into the selection of a model for a particular cost-effectiveness analysis should be explored. First, remember that the basic function of the model is to determine the value of the measure of effectiveness as an output when the controllable variables (i.e., those variables which are to be selected as a result of the analysis) are used as inputs. In the real world, we would set the controllable variables, let the uncontrollable variables (i.e., those set by higher authority or by the world outside the system itself) apply, and then observe the measure of effectiveness. In the model world, however, we must explicitly bring to bear the uncontrollable variables, and we apply only those that are particularly relevant to the question under consideration. Here we have a major design problem, and an area that is still very much an art: which of the uncontrollable variables are relevant--and many will inevitably have some relevance--and should be included in the model. In this sense, relevance of a variable relates to the sensitivity of the measure of effectiveness to the value of that variable over its reasonable range, and the interaction between that variable and the controllable variables. If a particular variable interacts negligibly with all the controllable variables under consideration, then that uncontrollable variable is not highly relevant to the analysis.

Another way of expressing this point is to require that the model be particularly sensitive to the differences between the alternatives being compared.

There will be the inevitable compromise between completeness and complexity, and between effort expended in modeling and savings achieved. There is clearly a point of diminishing returns in completeness of a model. The elaborateness of the model must also reflect the uncertainty of the input data available; if the only available data are crude, precise models are pointless.

Thus, after the appropriate measures of effectiveness have been defined, it becomes necessary to identify the controllable variables relevant to the decision under consideration. This, of course, depends strongly on the level in the decision-making hierarchy of the decision-maker being served. This level defines the scope of the model. For example, a decision of the Joint Chiefs of Staff regarding targeting in a general war situation would take as fixed the CEP of an ICBM and would use a model relating some value-destroyed measure of effectiveness to alternative targeting strategies; whereas a guidance-equipment manufacturer, who has CEP as a measure of effectiveness, must use a model relating CEP to the characteristics of alternative system designs with alternative gyros as controllable variables. Similarly, the missile manufacturer would take the guidance system behavior as input information and choose the missile design that gave the best total CEP, including the guidance system as a component. His controllable variables would then include guidance systems, and his model would relate guidance system performance to total CEP.

Having thus established the scope, the controllable variables, and the measure of effectiveness to be expressed as a function of these controllable variables, the model we choose depends on how well we know the relationship between the measure of effectiveness and the controllable variables. If this relationship is well known, especially where the relationship is a technological one, then we might try to develop an analytical model. Here, we might seek a system of differential equations which we can try to solve. The methods of calculus are ideally adapted to finding the maximum or minimum of a curve--and this offers a means for selecting the optimum among systems that vary in degree. Similarly, wherever we have two cost components, one of which is increasing with the value of the controllable variable and the other decreasing, we can use the calculus to find the optimum interval. In establishing an automobile lubrication interval, for instance, longer intervals lead to lower maintenance costs but to higher repair costs later in the vehicle's life. Calculus can find the optimum interval.

For more complex relationships, the calculus of variations and its many offspring is available for finding the optimum of a function subject to constraints, and even fairly complex functions can be handled in a variety of ways. For instance, we could find an optimum aircraft trajectory to reach a 20,000-ft altitude in a minimum time, or find the flight path of a commercial airliner to minimize fuel consumption subject to acceleration constraints. These techniques are particularly applicable to problems with relatively few variables and relatively few constraints, even though these may have fairly complicated but known relationships.

Where the relationship can be expressed in a linear form, even though there may be a very large number of controllable variables and a large number of constraints, the techniques of linear programming become applicable. Here, for instance, we can formulate the problem of minimizing the cost of shipping from many wholesale to many retail points, given constraints of capacities of the wholesale points and needs of the retail points, where the transportation costs are proportional to distance between the points. Many network problems can be treated as linear-programming problems, and optimum network configurations to minimize costs and meet capacity constraints (where capacity is proportional to cost) can be determined.

Game theory represents a class of models for developing an optimum strategy to be played against an opponent who likewise has a range of strategies open to him. This requires the payoffs to both sides to be fairly well known for any combination of strategies.

Queueing theory has been applied to a wide variety of situations that might be characterized as "service processes," in which a "server" services a succession of "customers," and there is an uneven match between the arrivals of customers and the availability of the server to provide service because of chance variations in either process. In any service process, we have a cost of serving and a cost of waiting, and any improvement to the service process reduces the waiting cost at the expense of the service cost. The problem, then, of finding the optimum service procedures can be solved using the methods of queueing theory. Here, the generality of the concepts can permit many kinds of processes. In addition to conventional super-market, banking, gas-station queues with which we are all familiar, we can consider a traffic intersection to be a server and the automobiles being served, a taxi stand represents the passengers as servers when the taxis queue up and the passengers as customers when the stand is empty.

Most importantly, in addition to special cases of these general models, one usually finds himself having to develop his own model. Here, all the power of mathematics is available, and probability theory becomes an essential part when stochastic processes enter the system's operation.

Where the system contains many components, and where each of them has a relationship to the others that is fairly well definable, or to entities processed by the system in time, then a simulation may be the most appropriate form of model. A good illustration here is a factory producing many products. The processing of an individual item on a single machine is a relatively straightforward process that can be described simply. The problem is one of describing the interaction of all the machines and the bottleneck that occurs at one particular machine because of breakdown of another, or because of an unanticipated change in the product mix. These possibilities can be examined readily by a simulation that examines each machine in succession. Simulation is probably the most popular technique now in use, and its use will undoubtedly increase with the growing availability of large computers to handle the large problems and the growing availability of modern simulation languages like Simscript. Furthermore, to develop an analytical model takes a degree of analytical skill, whereas anybody can write a simulation. Thus, the forces pushing in this direction are very strong, and we will see far more of it than we already have. It is an easy way out of the shortage of competent mathematicians, and supports our computer industry admirably.

Where the role of the human operator or decision-maker in a situation is important, then we would want to simulate him directly. For that case we would resort to gaming as a primary means of examining alternative systems. In the air traffic control example, for instance, we would test the effect of different control procedures on delay by using real controllers controlling simulated aircraft.

The operational exercise is most useful for a final comparison of real operating systems to introduce the field hazards and unexpected eventualities and to provide basic input data to the other models. The effort at the Combat Developments Evaluation Center, for instance, performs this role for the Army to a major degree.

VI. SUMMARY

In summary, we have shown that the model provides a means for determining the measure of effectiveness as a function of the controllable variables in terms of the relevant uncontrollable variables, and we have characterized the various classes of models as operational exercises, gaming, simulation and analytical, in increasing order of abstraction. These are applicable to different classes of problems: the exercise for last-stage testing of systems, gaming where human decision plays an important role, simulation for complex problems which are internally describable in complete detail, and analytical models to make use of the power of mathematics, especially where we are searching for an optimum.