AD 686757

RESEARCH ANALYSIS CORPORATION

Proposal Generation and Evaluation Methods in Research and Exploratory Development

RAC

Reproduced by the
CLEARINGHOUSE
for Federal Scientific & Technical
Information Springfield Va 22151

DISTRIBUTION STATEMENT
This document has been approved for public release and sale; its distribution is unlimited.

ECONOMICS AND COSTING DEPARTMENT
RAC PAPER RAC-P-11
Published November 1965

DISTRIBUTION STATEMENT
This document has been approved for public release and sale; its distribution is unlimited.

Proposal Generation and Evaluation Methods in Research and Exploratory Development

by
Bernard Sobin
Arnold Proschan

RESEARCH ANALYSIS CORPORATION
McLEAN, VIRGINIA

This material is not to be reproduced for publication, sale, or general distribution. It will be included in RESEARCH PROGRAM EFFECTIVENESS to be published and copyrighted by Gordon and Breach, Science Publishers, Inc., 150 Fifth Avenue, New York, N.Y. 10011.

Received for Publication 23 August 1965
Published November 1965
by
RESEARCH ANALYSIS CORPORATION
McLean, Virginia

FOREWORD

This paper was prepared for presentation at the Second Conference on Research Program Effectiveness, Office of Naval Research, Washington, D. C., 27–29 Jul 65.

CONTENTS

Introduction 3

Proposal Generation 4

Characteristics of the Proposal-Generation Process—Proposal-Generation Search Proceeding from Materiel Objectives—Proposal Generation Proceeding from the Technological Base—Role of Central Management

Evaluation 10

Characteristics of Project Proposal Value—Use of Subjective Judgment in Comparative Evaluation—Multiple Rating-Scale Formulas—Dealing with Ordinality of Subjective-Value Comparisons—Definition of Project Proposals

References 25

References Cited—Additional References

Figures

1. Elements of Project-Proposal Value 11
2. Value as a Function of Order Number for Proposals at an Army Laboratory 21
3. Value as a Function of Order Number for Hypothetical Data 22

Table

1. Worksheet for Display of Value as a Function of Order Number at an Army Laboratory 20

Proposal Generation and Evaluation Methods in Research and Exploratory Development

ABBREVIATIONS

CAMD	combat assault mobility device
DOD	Department of Defense
QMDO	qualitative materiel development objective
R&D	research and development
RDTE	research, development, test, and evaluation
USAMC	US Army Materiel Command

INTRODUCTION

The process of determining a preferred program of research and development (R&D) in science and technology may be described as consisting of three steps:

(1) Generation of a worthwhile list of projects.

(2) Characterization of projects according to their value in serving the objectives of the program.

(3) Application of a mathematical model for maximization of value subject to resource constraints.

This paper first treats the first step, confining attention largely to the exploratory-development stage and then treats the second step, the characterization of projects according to their value. The third step is not treated in this paper except for some comments on the reduction in resource allocation efficiency when the evaluation is incomplete.

This paper is based on a RAC investigation of ways to improve Army planning for research and exploratory development. AR 705-5[1] defines these categories of the Department of Defense (DOD) research, development, test, and evaluation (RDTE) program as follows:

Research. Includes all effort directed toward increased knowledge of natural phenomena and environment and efforts directed toward solving problems in the physical, behavioral, and social sciences that have no clear, direct military application . . . .

Exploratory Development. Includes all effort directed toward solving specific military problems short of major military developments projects. It may vary from fairly fundamental applied research to quite sophisticated breadboard hardware, study, programming, and planning efforts. It would thus include studies, investigations, and minor development effort. The dominant characteristic is that the effort is pointed toward specific military problem areas with a view toward developing and evaluating the feasibility and practicability of proposed solutions and determining their parameters . . . .

PROPOSAL GENERATION

CHARACTERISTICS OF THE PROPOSAL-GENERATION PROCESS

At last year's conference on research program effectiveness, David B. Hertz pointed out the importance of proposal generation in the research-planning process. The number and quality of the proposals from which a program is drawn obviously have a major bearing on potential achievement. This aspect of program planning is therefore appropriately considered before the problems that decision makers face in choosing among alternatives. It should be recognized, nevertheless, that proposal generation and selection are interrelated, since one cannot search and design without at least some implicit ideas of relative value and resource implications, and, conversely, a good resource-allocation methodology may well serve as an influence toward improved design of candidate projects for application of effort. Additionally, in a large organization a proposal may pass through several stages where proposal generation and selection concerns alternately predominate.

Proposal-generation procedures will of course be influenced by the objectives of the effort. The exploratory-development category of the DOD RDTE appropriation includes effort directed toward expanding technological knowledge and developing materials, components, devices, and subsystems that are hoped to have military relevance. Examples of successful exploratory developments are techniques for increased reliability of certain communications equipment, improved armor made of ceramic-faced material, a technique for the conversion of a conventional airborne radar to improve its utility for reconnaissance, and fibers having high stiffness-to-weight ratios for reinforcement of structural composites.[2] Better than these examples for an understanding of the kinds of exploratory-development projects that are accepted would be a statistical analysis of exploratory-development activity by such classification schemes as degree of hardware orientation, degree of application, anticipated time for final application, and degree of advancement sought. Such an analysis is part of a study being performed by RAC for the US Army Materiel Command (USAMC).

The objectives of exploratory development previously described obviously derive from the more ultimate objective of R&D—the development of new or improved materiel for the use of the military forces. The general problem of developing candidates for effort in exploratory development may be approached by attempting to define the ultimate materiel items desired and then working back from them to possible technological means to achieve them. Illustrative

of this approach is the process common in the military services in which user requirements are defined as a guide to the development organization.

The generation of candidates for effort in exploratory development may proceed, not from ultimate materiel goals, but from more intermediate goals of achieving technological improvements. There may be only the haziest of ideas in the program planner's mind as to the specific materiel items in which the technological advances may ultimately be exploited.

When the materiel item is perceived at the proposal stage, the search procedure preceding the proposal formulation could have started with a materiel problem leading to a search for technological approaches; or it could have started with an initial interest in an intriguing bit of new scientific or technological knowledge followed by a search for a materiel application. When the materiel item is not perceived, the search emphasis must be on exploiting new scientific or technological ideas or on a general broadening of the technological base. Approaches to search from the materiel ends and from the technological means will be discussed separately; they are not, however, to be considered as independent.

PROPOSAL GENERATION PROCEEDING FROM MATERIEL OBJECTIVES

The approach to the generation of candidates for effort from the direction of materiel objectives may be made by the user, the developer, or by some combination of user and developer.

A formal instrument for user expression of requirements in the Army is the qualitative materiel development objective (QMDO), defined in AR 705-5[1] as "a statement of a Department of the Army military need for developing new materiel, the feasibility or special definition of which cannot be determined sufficiently to permit establishing a qualitative materiel requirement" (the latter is the medium for stating needs for materiel items believed to be feasible). The titles of the QMDOs indicate their nature, e.g., random-access discrete-address systems, an active armor-protection device, and expendable small craft. QMDOs are promulgated in a Combat Development Objectives Guide, which serves the purpose of providing guidance for the development of future operational concepts, organizations, and materiel; it correlates combat development activities with the R&D program.

Recently the Army has moved toward securing more definitive QMDOs and has sought to upgrade the process of planning for the attainment of the objectives of QMDOs. The elements of QMDO planning, as prescribed in a recent revision of AR 705-5, include alternative approaches and a definition of the anticipated barriers to success in each case and task networks for the approaches considered feasible. More time will evidently be required to secure extensive accomplishment of this kind of planning.

A recent experiment of the USAMC provides good insight into how developer-derived concepts of desired materiel items can serve as a basis for the generation of candidates for effort in exploratory development. Over the last 2 years USAMC has conducted a significant and productive experiment in long-range technical planning, of which an important element was the determination of tasks required to be accomplished to achieve a designated end-item

of equipment with sufficiently advanced characteristics as not to be within the current state of the art. A specific system objective, a combat assault mobility device (CAMD), was postulated and analyzed in detail as a means of getting at the pacing technology and formulating specific exploratory-development tasks.

Based on USAMC's technical objective guide, critical performance parameters and major functional requirements were determined. Further analysis identified the specific problem areas involved. These were logically interrelated in network form to show the various efforts required, such as analysis studies, experiments, and component developments. Task descriptions were developed to cover the problems, approaches to be taken, goals to be achieved, probabilities of success, and manpower and funding required.

Exploratory-development tasks defined as a result of such a planning process are not undertaken for the achievement of a specific defined system, i.e., they are not part of a weapon-system program. Rather they are efforts to push back technological frontiers in areas where there is some reason to believe the payoffs would be substantial. The weapon system chosen in the USAMC experiment, CAMD, is intended to be representative of a class of possible systems, not a firm plan for a particular one. Not only because of the uncertainties involved in the innovation process but because of possible changes in the strategic situation or enemy capabilities, too sharp a focus at the exploratory-development stage on the ultimate equipment items desired may have adverse effects. Such a focus may overly limit the work in the technological areas of primary concern and overcommit efforts to problems that subsequently turn out to be somewhat irrelevant.

Nevertheless, some of the participants in the experiment whose duties normally were to develop specific system designs found it difficult to fully recognize that their objective was to determine barrier problems for attack rather than to plan complete systems and supporting programs.

There was also a problem in the workload involved in comprehensive networking of a major system. Had the scope of the experiment extended to alternative concepts of the CAMD, an even more extensive workload would have been involved. A process more selective than comprehensive networking of a system is needed for convergence on barrier problems.

The experiment provided valuable experience in the determination of the critical tasks to be accomplished for the realization of an advanced materiel item. Such experience should be of direct benefit in improving the planning (including networking) of exploratory-development effort supporting QMDOs.

An additional benefit of the experiment was the improvement of interlaboratory planning by getting planners accustomed to working together, a particularly significant achievement in a command encompassing what had previously been separate technical services.

As an alternative to user or developer expression of desired materiel items as a basis for the generation of candidates for effort in exploratory development, a combined user-developer approach may be chosen. Such an approach would involve a free association of force planners and developers, perhaps in the form of a mixed team working closely together to achieve a unification of force planning and technological considerations. Various concepts for equipment would be hypothesized and then evaluated in the context of a future force structure and situation to gage potential contributions.

The hypothetical concepts would not be "given" to the team; they would be derived by the team from forecasts and state-of-the-art analyses and from studies of functional requirements of future forces. Obviously a strictly sequential process is not involved here but rather an iterative process in which anticipations of force requirements and technology interact. From this force-planning exercise, preferred hypothetical concepts would be derived to serve as a focus for long-range planning.

For the preferred hypothetical concepts, there could be a selective networking of key problem areas, thus sequencing effort in time toward concentrating exploratory-development effort on the pivotal problems. The importance of selectivity of effort is highlighted by Dr. Suits[3] in this passage, which follows his description of the case history of the high-power-vacuum-circuit interrupter:

> A key technical problem was identified early in the project. Success or failure turned principally on the question of maintenance of vacuum. Unless a positive solution to this problem was found, the concept was impossible to realize. There were many other important and interesting problems, but the temptation to solve the peripheral problems was resisted in favor of concentration on the pivotal problem. This key problem was attacked repeatedly as new approaches were devised, or became available.... If there is a pivotal problem, and if you can identify it, economy of effort favors concentration on the solution of this essential problem first.

PROPOSAL GENERATION PROCEEDING FROM THE TECHNOLOGICAL BASE

Preceding paragraphs have described proposal-generation procedures emanating from concepts of desirable materiel items, possibly those developed from a consideration of such items in the context of a force structure. Much of the work at the exploratory-development stage derives from objectives not directly related to overcoming the barriers to fairly tangibly conceived materiel items, e.g., improved characteristics, techniques, processes, measuring devices, and experimental procedures. Such a situation obtains because of a common reading of experience by both researchers and managers that the serving of these proximate objectives is an effective means for achieving in time the ultimate objectives, i.e., that the availability of more advanced technology will lead to exploitation in more advanced materiel. This thought is reflected in Standifer:[4]

> The other avenue of applied research has its origin in the results of basic research and the intriguing promises they hold out. Although an end use is not always immediately discernible, there are attractive implications of an eventual enhanced technology—and there are few technologies that have no value to the Air Force. Experience has shown that, inevitably, good research will find or develop its own application. Moreover it is this kind of inquisitive applied research which, as much as anything else, produces those unexpected breakthroughs, breakthroughs from which entirely new possibilities for advanced systems are frequently conceived.

At any time there are numerous well-recognized technical problems to be solved stemming from deficiencies of operational equipment, difficulties in completing developments on which considerable effort has already been expended, known causes of failure for abandoned projects, and impediments to advances in performance characteristics. As new knowledge or

technology becomes available, it is screened by many researchers against the endless list of problems to determine where new or additional effort offers promise of achievement. This process of matching takes place ceaselessly. Creativity consists in seeing possibilities for paths and combinations that do not readily suggest themselves. Such a process does not lend itself well to centralization; hence broad diffusion of knowledge of both military materiel problems and new knowledge and technology is a key factor in promoting advances depending on the recognition of new possibilities.

Stimulation of the matching of problems and possibilities can be provided through emphasis on state-of-the-art forecasting. Such forecasting may well foster some preliminary planning to achieve the new technology. Additionally, knowledge of new technology expected to be available will suggest to some the value of other kinds of technology that would be feasible and valuable. Broad laboratory participation is desirable with forecasts of greater depth in selected areas through research conferences and industry studies. Industry participation may also stimulate industrial initiative on technical proposals for exploratory-development work.

When new knowledge or technological accomplishments are recognized to be of outstanding importance, a more directed effort toward exploitation can be undertaken. Special effort may be allocated to the analysis of applications of recent technological advances. Such an approach is illustrated in a program of the Advanced Planning Division of the Naval Analysis Group of the Office of Naval Research, which involves study of such concepts as (a) potential utility of all-magnetic logic and signal-switching techniques to Naval systems, (b) utility of pure fluid control and amplification devices, and (c) utilization of associative memory concepts in computer, data-processing, and information-retrieval systems.

It is understandably difficult to develop statements of objectives that are meaningful for proposal generation proceeding from the technological base; a multiplicity of intermediate objectives are necessarily involved. Perhaps this is why R&D planning documents are often limited in their utility. Such plans may convey an impression of vagueness or generality; they may lack overall integration and relative emphasis. In actuality, program formulation must in large measure be accomplished directly by those most closely associated technically with the various fields. A key role in exploratory-development program formulation is, as a consequence, indicated for laboratory staff and management.

ROLE OF CENTRAL MANAGEMENT

In emphasizing the role of laboratory management in program formulation for advancing the technological base it is not intended to minimize the contribution that central management can make. Central management can indicate areas for additional emphasis based on the evaluation of such items as enemy capabilities, operational difficulties, problems with newly developed equipment, forecasts, reports of reviews by panels of experts, contractor proposals, capabilities of the R&D organization, and high cost of existing capabilities or of proposed improvements. It is an essential function of the highest

technical management to synthesize these and develop ideas on areas in which broader effort is needed. Response to such pressures can come from the operating technical levels in the form of specific proposals for augmented effort.

Central management can make more general contributions in encouraging imaginative thinking, protecting exploratory development against excessive diversion by near-term problems, stabilizing military management, promoting military planner and researcher interchanges, and other such management emphases. These, however, affect not only the process of program formulation but the entire R&D process.

EVALUATION

CHARACTERISTICS OF PROJECT PROPOSAL VALUE

There are, of course, many ways to classify the elements of project-proposal value. Figure 1 is a diagram of what is believed to be a useful classification. The major categories are (a) expected value of the knowledge that will appear in the project report or in other communicable form and (b) values in the research process itself, excluding communicable knowledge generated by the project. Since the first category is an expected value, it consists of a sum of products of achievement combination values and their probabilities; or, if the combinations are of continuous variables, it is a multiple integral of products of the value and probability functions of an infinite variety of combinations. The note on complementarity in the figure is intended to highlight the fact that, in the multiplicative relation between individual-knowledge item values and their associated probabilities, neither can make any contribution to value unless associated with a positive value of the other. The figure also notes some of the kinds of elements or indicators of the probabilities of knowledge achievement and of the values of knowledge items and the research effort. A few comments may be useful on the significance of this somewhat generalized structure of project-proposal value to the case of military research and exploratory development.

An item of knowledge appearing in a project report, however basic that knowledge, derives its value for military purposes from its expected effect on weapon-system technology over some extended period of time. Such technology improvement can conceivably involve no change in capability over this period of time except a reduction in cost of that capability. Then the cost saving is clearly an index of the value of the technology. Generally, however, the new technology will best be used to increase military capabilities to levels at which the cost may be either higher or lower than in the absence of the new technology. Then the index of value is multidimensional, involving cost changes, one or more kinds of capability changes, and different dates of capability. It is assumed that costs incurred at different dates can be summed after suitable discounting of the later costs.

Even when the project report value is multidimensional, however, it is sometimes possible to obtain useful upper and lower bounds to an aggregate representation of value measured in units of cost alone. A lower limit on the money value of an increase in capability is the saving that would occur if the new technology were substituted for the old in achievement of the lower capability. An upper limit is the saving that would occur if the new technology were

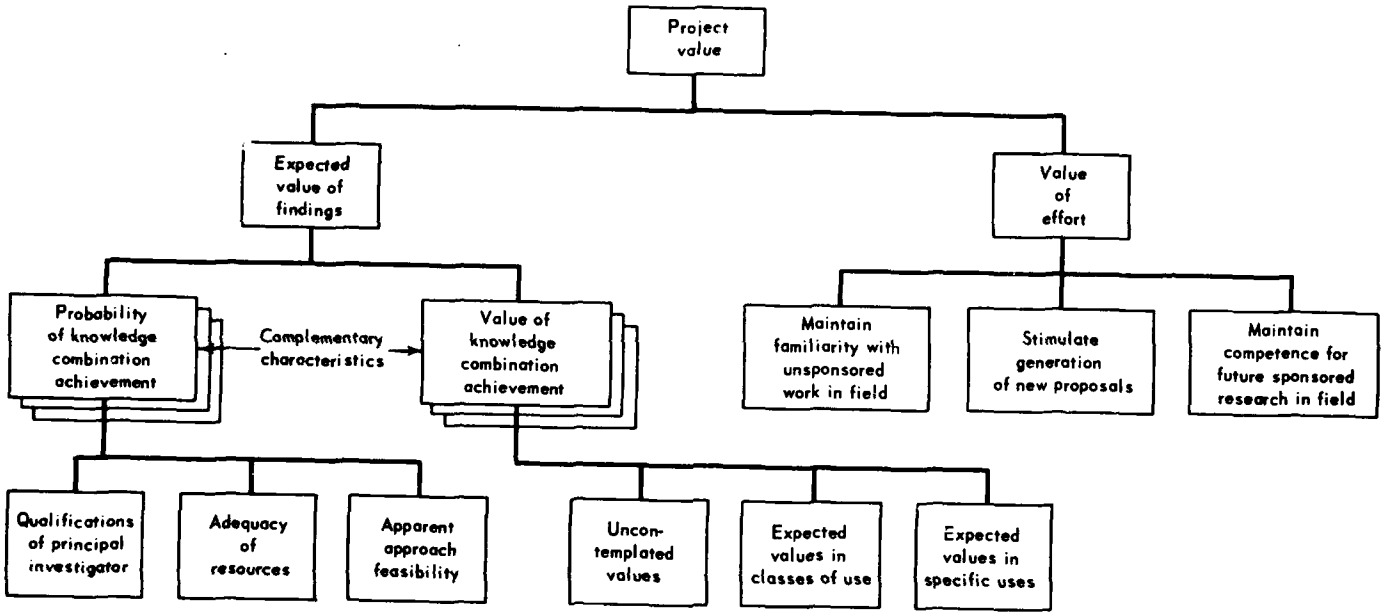

Fig. 1—Elements of Project-Proposal Value

substituted for the old in achievement of the higher capability. The first saving is a lower limit for new-technology value because it assumes a less than optimal exploitation of the new technology. The second saving is an upper limit for new-technology value because it assumes a less than optimal way of dealing with the absence of the new technology. An earlier discussion of these upper and lower limits to value appears in Berman.[5]

In exploratory development, there can be projects each of which is intended to determine the feasibility of the key technologies underlying a particular new kind of weapon system that has fairly identifiable characteristics. Then it may actually be practical to attempt rough calculations of the money value of the project report. Most exploratory developments, however, involve technologies applicable to classes of subsystems across not necessarily identifiable complete weapon systems. Research projects, by definition, are not tied to identifiable complete weapon systems, although the vagaries of accounting decisions might allow a few exceptions. In cases where weapon systems are not identifiable, explicit calculations of value in this way become impractical.

Where there are foreseeable types of military applications the ability to specify the complete weapon systems that would use the technology helps, of course, in the estimation of value of the technology. However, exclusiveness of the application to a particular weapon system of detailed configuration and use tends to reduce the value, particularly when the weapon system is far in the future. The further one goes into the future, the less reliable are forecasts of the technology that will be available and of the enemy threats that new weapon systems must counter. Any presently planned weapon system that is optimized for a future technology and threat becomes less valuable as less is known of such future circumstances; and any current R&D project that serves exclusively to support such a weapon system suffers from the same weakness.

The probabilities associated with possible combinations of knowledge item achievement are clearly not subject to explicit calculation. Skilled project evaluators, however, have rough indicators of the likelihood of high or low levels of achievement, as indicated in Fig. 1. The most important of these indicators, particularly in research, is frequently considered to be an appraisal of the qualifications of the principal investigator, which, in turn, are indicated by his record on previous projects, formal education, publications, and the opinions of experts who have had close contacts with him. The adequacy of availability resources and the apparent plausibility of the investigator's planned approach to the problem are also clearly relevant; but, if the principal investigator is regarded very highly, his opinions on these points will carry important weight with the evaluator.

The qualifications of the principal investigator may also serve an evaluator as an indicator of the importance of the types of knowledge to be sought, particularly if the evaluator does not have an independent idea of the importance of the knowledge or if he has no clear idea of exactly what knowledge is sought. It should be noted, however, that the investigator's qualifications that make for high probability of success in finding answers to specified questions need not be correlated closely with judgment on the importance of the questions.

Figure 1 lists three kinds of value that are associated with doing project work, considered independently of the value to the military of reports and other communicable outcomes. (Of course, there may be a high correlation between

values of project findings and the values of project activity considered here. There is no independence in this statistical sense.) The first of these additional kinds of value is based on the view that a scientist has better access to the work of other scientists if they are also interested in his work and if he is respected as a contributing member of the scientific community. The second is based on the fact that a laboratory usually cannot keep a specialist for later assignments unless he is supplied with current work he considers worthwhile. In addition the research may maintain and advance his competence more effectively than would only keeping abreast of the literature. The third is based on the view that creativeness is increased by research activity. All have in common the fact that they maintain or advance the capabilities of research personnel.

No opinion is offered here on the relative importance of these research-activity values compared with values of the items of knowledge referred to on the left side of Fig. 1. Whatever their importance a point to note is that they are not unique to particular projects. Each kind of value can be achieved by any of a class of projects. Moreover, successive increments of projects serving a class of values will eventually reach a point of sharply diminishing returns. This suggests that these values should be disregarded in any evaluation of individual proposals that precedes decisions on resource allocation.

USE OF SUBJECTIVE JUDGMENT IN COMPARATIVE EVALUATION

The term "subjective judgment" as used here may be defined as a way of arriving at a conclusion by a mental process that is not well enough understood by the practitioner to enable him to demonstrate to others the relation between his conclusion and premises. Operations analysts and the general scientific community may properly be uneasy about the use of such conclusions since they have no way of checking either all the premises or the logic. On the other hand, situations exist in which the kinds of data needed to support logical conclusions either do not exist or exist subject to such great measurement error that negligible confidence can be placed in the conclusions arrived at by objective analysis. A review of the elements of project-proposal value makes it apparent that the evaluation of research and exploratory-development proposals is an example of such a case. It is therefore important to maximize the quality of subjective judgment and use it where it may be most effective.

An obviously necessary condition of good judgment is understanding by practitioners of the problem to be solved, which may in other fields involve judgments on matters as diverse as estimating a distance from an observer to some other point and determining which of many alternative courses of action is more likely to lead to victory from a given position in a chess game. (A number of large electronic computers have been programmed to play chess using all known principles of chess strategy and tactics, but a good chess player can regularly win against any of these computers.) In the case of R&D project-proposal evaluation, there is little hope of good subjective evaluation, no matter how eminent the practitioners may be in particular academic fields, unless they understand clearly management's goals in sponsoring R&D activity and are willing to value proposals solely according to the degree that the proposals will promote those goals.

A second characteristic of good subjective judgment is past experience of practitioners with a number of combinations of the kind of thing to be estimated and observable phenomena that might be useful indicators. In subjective distance measurement, each combination might be a distance actually measured and such phenomena as apparent size of objects of known size, distinctness of the view, and some stereoscopic effect of observations with two eyes. In the case of the chess position, each combination might be an actual game outcome from some position and such immediately observable characteristics of the position as the amounts of each kind of material held by each player and certain technical characteristics of the position considered relevant by chess players. In R&D project-proposal evaluation the relevant experience would be with combinations each of which would consist of an actually observed indication of project value and such associated proposal characteristics as training of principal investigator, area to be investigated, and kinds of resources to be used.

If these experience data are bases for good judgments, it may be presumed that the data are processed in some logically sound analytical model before the judgments are reached. An intuitive analysis along the lines of statistical-correlation analysis appears plausible.

When project-proposal evaluators lack a clear understanding of the values to be maximized, they can be indoctrinated rapidly (along the lines of Fig. 1 or some elaboration of it) with the objectives of the organization and the effects their decisions might have on those objectives. Supplying missing kinds of experience is bound to be a slower process. The evaluators might draw on what experience they have, to determine tentatively the kinds of observable characteristics of project proposals likely to be relevant to correct concepts of value, and they should immediately begin to accumulate data and experience to provide bases for continuous revision of hypotheses on how to compare proposal values. The growth of judgment quality may be slow and uncertain because of time lags before the effects of proposal acceptance can be observed and because of changes over time in the real significance of different observable criteria of value. Such changes of significance of past information do not occur in distance measurement or in selection of chess strategy.

The improvement of subjective judgment should be associated with efficient selection of the part of a problem to which subjective judgment should be applied. In allocation of resources to R&D projects the judgment can conceivably be applied to (a) the resource allocation, (b) the overall evaluation of projects, or (c) the ratings of projects with respect to specific characteristics. The following sections covering specific suggestions for evaluation improvement may also be considered discussions of the proper scope of subjective judgment in the broad problem of resource allocation.

MULTIPLE RATING-SCALE FORMULAS

Comments in preceding sections on multidimensionality of value and lack of data for explicit calculation of either aggregate value or elements of value should come as no shock to those who have faced evaluation problems. The statements on the need for experience and knowledge as a basis for good

subjective judgment are also familiar ideas. Operations analysts, building on ideas similar to those expressed here, have frequently recommended what will be described in this paper as multiple-rating formulas.

A multiple-rating formula aggregates to a single score the ratings of a proposal according to a number of different criteria. The ratings are the variables in the formula. The rating scales, the type of formula, and the numerical constants of the formula are determined in advance for all proposals to be reviewed, and the proposals differ with respect to the ratings assigned.

Most common multiple rating formulas are linear, i.e., the formula score is a weighted sum of ratings. Implied in a linear formula is the principle that the value contributed by each rating is independent of the other ratings. Nearly all nonlinear formulas in the literature are multiplicative. Here, total value is the score reached when all ratings are multiplied against one another.* In such a formula, perfect complementarity of ratings exists in the sense that a high rating in any respect is of greater significance if the other ratings are also high; and a zero rating in any single respect precludes any aggregate value, regardless of the levels of other ratings. Rarely, formulas will have mixtures of linear and nonlinear parts, and sometimes a formula will be supplemented with minimum levels of certain ratings as side conditions for proposal acceptance.

The criteria according to which proposals are rated in such schemes are defined so that there should be prospects of making either accurate objective estimates or expert subjective judgments, and, for any criterion that requires subjective judgment, an attempt is made to have that judgment applied by one or more experts on that criterion. A principal strength of multiple-rating formulas is therefore the quality of the subjective judgment that can be brought to bear on the ratings.

The choice of criteria, the scales for rating, and the formula for aggregating ratings also involve subjective judgment, although some objectively determined restrictions can usefully be placed on the judgments. How expert can these judgments be? For an indication of an answer, one may start with an appraisal of the accuracy of a formula consistent with an assumed-value classification similar to that in the left major branch of Fig. 1.

The formula inventor might agree that the three boxes at the bottom of the figure are good definitions of criteria for estimating probabilities of various degrees of success in achieving knowledge goals of the project. He would then need to decide whether the three elements in the probability (qualifications of the principal investigator, adequacy of resources, and apparent approach feasibility) are independent of each other (simply additive in their effects) or complementary (multiplicative or perhaps having more complicated interactions). Suppose he should decide on independence—although it is obvious that there must be some complementarity, since the most generously supported project with the most promising technical approach would have little hope of success if the principal investigator should be a known incompetent. (For a discussion of independence tests by gamble preferences, see Fishburn.[6]) He would need to determine the proper scales and standard weights for estimates of these criteria.

*When a rating has a scale such that the higher the number, the less desirable it is, the multiplication in the formula may refer to the reciprocal (e.g., cost) or to the complement (e.g., probability of failure).

Even more difficult problems would arise in the valuation of the knowledge if successfully achieved. Assume that there are some known military uses, values of which are not capable of being estimated directly. The formula inventor might then decide that military experts should rate the aggregate of these uses with respect to one or more of such criteria as importance of the missions involved, adequacy of current capability with respect to those missions, and degree of improvement in mission capability for given cost. But then he would have to determine how these ratings might be aggregated in ways that took proper account of complementarities, substitutabilities, and relative importances.

In dealing with uncontemplated values of the knowledge and with values in classes of uses, the formula inventor might fall back on criteria of "scientific worth," or "degree of technological advance," which are popular terms among evaluators. He might also use a little additional weight for the qualifications of the principal investigator, on the theory that the kind of project that interests a highly qualified man is likely to deal with a question whose significance will become apparent later if it is not apparent already. In such situations the scales and formula parameters are likely to be particularly arbitrary. Not only will individual projects differ with respect to the significance of the criteria, but there is also likely to be some difficulty in finding good operational definitions of the criteria.

Actually the typical inventor of formulas and rating scales is not as systematic as our hypothetical one in tying the particular criteria and the types of relationship among criteria to a specific concept of value such as that of Fig. 1. On the contrary, many of the additive formulas in the literature have independent terms for criteria that clearly represent elements of value that are complementary (e.g., criteria of success probability and of value if successful, or criteria of mission importance and of effectiveness in advancing missions), and purely multiplicative formulas have factors representing value elements that are clearly independent (e.g., usefulness for two or more independent purposes). But, on the assumption that these obvious weaknesses or multiple rating-scale formulas are eliminated, what is their potential?

The discussion of how a hypothetical formula might be developed indicates that it is not easy to develop a "best" formula. The separation of formula invention from the rating of specific projects according to formula criteria has the merit that each kind of work can be done by a specialist in that field. It is not clear, however, that even a best single formula is an accurate instrument, since the best formula for one project or group of projects may conceivably be far from the best for another project; e.g., the qualifications of the principal investigator could be of critical importance for some projects but not for others. In fact the formula may be poor even for the project to which it applies best. This last point could be condoned more easily if the formula were indeed the best way of organizing all the information that might be available about a project. But the formula covers only those items of information that can be defined for rating purposes across a great many projects unless, which has not been previously mentioned, the formula inventor decides to make a place in the formula for the variable "miscellaneous evaluator impressions." For example, an evaluator might be very interested in a project because of prospects of its having findings that would set an upper limit to a class of

performance of certain kinds of weapon systems and therefore serve as a means of avoiding waste of research effort. Such formal criteria as scientific merit, advance of technological base, and importance of military mission might not give this consideration the importance it deserves. The division of labor between formula inventors and raters has the general disadvantage of failing to use the best available data and formula for each individual project in favor of using the best that can be uniformly applied across all projects in competition.

This is a serious weakness although perhaps not serious enough to outlaw the use of such formulas entirely. It may be possible to afford considerable error in the choice of rating scales, types of formula, and formula parameters, provided that the bulk of the true elements of value, whether or not specifically known, can be assumed to be positively correlated with a number of formula variables statistically large enough so that many have significant weight in the formula. Then even if the scales and formula parameters are subject to large errors those errors will not all have effects in the same direction. The net value error after offsets among formula errors may frequently be expected to remain small enough for the calculated values to be useful.

Whether a particular formula has the statistical qualities needed to overcome any of its weaknesses in the use of available information on types of criteria and their interactions, in choice of formula parameters, and in determination of rating scales is a matter to be decided by the principal user of the formula. If, after proper indoctrination in the strengths and weaknesses of this formula, he thinks it helps him, it probably does. It is probably harmful if he does not like it.

The evaluator also has the alternative of basing his appraisal partly on a set of expert ratings with respect to any criteria he considers suitable but without using any standard aggregation formula. Then he receives the benefit of uniformly applied expert judgment on specific points but retains flexibility on interpretation of the significance of those ratings for each proposal at issue.

DEALING WITH ORDINALITY OF SUBJECTIVE-VALUE COMPARISONS

When subjective judgments of value make no use of such devices as multiple-rating formulas it must be clear that the value judgments as expressed can only be ordinal, although ways exist (including one suggested later in this paper) for an analyst with data on a number of ordinal judgments to make some cardinal inferences. A multiple-rating formula yields for each project proposal a numerical index of value, but if the final evaluator uses the score only as one of other important inputs to his judgment, as suggested in the preceding section, the result of the evaluation is still ordinal.

Estimates of resource requirements are cardinal. One may estimate the percentage by which the cost of one project may be expected to exceed the cost of another project. But this cardinal information about resource requirements is of limited objective use if the costlier project happens also to be a more valuable one. If a choice must be made between the two projects there is no obvious way to determine objectively whether the higher cost is justified by a correspondingly higher value.

There are three approaches to the problem of making the best use of ordinal judgments on value and cardinal estimates of resource requirements in an objective resource-allocation procedure. The first is the use of mathematical-dominance arguments by means of which it can be demonstrated that certain of the proposals should be accepted or rejected, but unfortunately this assignment can be made with respect to only some of the proposals. The second and third approaches both involve further subjective judgment to supplement the original determination of value inequalities between single proposals. The further judgments are of value inequalities between small combinations of proposals. The second approach makes them as part of the resource-allocation procedure. The third makes them as part of a statistical model for conversion of nearly all the original ordinal information about value into relative values. Then the latter allows the use of standard kinds of programming models for the actual resource allocation.

When the resource-allocation model is to use only the basic ordinal information on the value of individual proposals it can use whichever of two decision rules is applicable: (a) accept a proposal if available resources are sufficient to include that proposal plus all others except those that have both lower value and higher cost; or (b) reject a proposal if available resources are not sufficient to include that proposal plus all others that have both higher value and lower cost.

These decision rules have the advantage of requiring no more information than is already provided in the cardinal information on resource requirements and the original ordinal information on proposal value. Moreover the conclusions can be demonstrated as necessarily implied by the premises, without any statistical errors. Unfortunately they will generally not be sufficient to fully determine a research program. After completion of such analysis as far as it can go there will still be a pool of proposals from which program choices can be made without exceeding resource restrictions.

The technique that is part of a resource-allocation procedure starts with a basic program resulting from a procedure that is in frequent use and progressively improves on that program by sequential introduction of additional preference information about combinations of proposals. The basic program is that associated with ranking all proposals in order of value without regard to cost and accepting proposals for the program in the same order until a budget restriction is reached. The details of the RAC procedure need not be described here. The essential idea is that each of the items in the tentatively included set of proposals (originally the basic program just described) is examined to see whether any pair of items in the tentatively excluded set, although lower in individual value than the item in the tentatively included set, has a higher combined value at no greater cost. When one or more pairs of items in the excluded set meet this condition, one of those pairs—which pair is one of the details not discussed here—is substituted for the single item in the included set. The procedure can be continued, if desired, with comparisons of kinds of combination pairs other than one item from the included set and two from the excluded set.

A few numerical experiments were made at RAC with a detailed procedure using the foregoing principles. In these experiments, random distributions of costs and values were used. The random numbers of value were, of

course, cardinal, but the procedure limited use of the cardinal information to determining inequalities between combinations of proposals and to final appraising of experimental results. The experiments were tried with and without correlation between cost and value, but not with all the kinds of value and cost distributions that were plausible. The experiments showed that, over the basic program, significant improvements in total value for any given budget limit were possible without the necessity of going to inequalities between combinations more complex to evaluate than comparing one proposal in the included set with two proposals in the excluded set. Nevertheless further work on the procedure was deferred because an alternative procedure appeared capable of giving better results with less effort and requiring no greater ranking capability on the part of evaluators.

It is expected that details of this procedure and of the practical problems appearing in Army laboratory experiments using it under RAC guidance will be published after more work has been completed. But the essential features, enough for a preliminary appraisal of the procedure, can be described here. Other procedures are described in Churchman and Ackoff,[7] Peck and Scherer,[8] Thurstone and Jones,[9] and Fishburn.[10]

The procedure starts with a ranking of projects in order of value from lowest (with order number 1) to highest. A graph is prepared with an arbitrary logarithmic scale of proposal values and an arithmetic scale of order numbers. The first plotted point is an arbitrary value for one of the projects with an order number near but not at the top. Because the number of highest ranking projects is small in relation to the value variation among them, the following procedure for estimating values of lower-ranking proposals would be too inaccurate in the immediate neighborhood of the top. The formula for best starting rank in different situations has not yet been determined, but the 90th percentile would probably be satisfactory in most situations. All other plotted values are to be one-half, one-quarter, one-eighth, etc., of the arbitrary one.

The key part of the procedure, an iterative one, is finding a proposal with half the value assigned to a proposal with previously plotted value. This is achieved by first finding the pair of proposals separated by two order numbers with aggregate value most nearly equal to that of the previously valued proposal and then assuming that the proposal with order number between those of the pair has half the value of the pair and therefore half the value of the previously valued proposal.

In a simple version of the procedure, this order number with half of the previously plotted value can then be the next plotted point, with the first plotted point (for a high-ranking proposal) having an arbitrary value. A chain of such points can be calculated, and then interpolations for values of order numbers between those plotted can be made.

A more refined RAC procedure, however, reduces the sensitivity of the slope of an interpolation line to the particular proposal at the upper end of the line. A more typical slope is estimated by using an average of three estimates. The first of these estimates is the number of order numbers for a halving of value from that of the whole order number closest to that of the last previously plotted point. The other two are numbers of order numbers for halving of value from the proposals ranked immediately above and immediately below that used for the first estimate. The new plotted point is then that of a value half that of the last previously plotted point and of an order number, perhaps

fractional, that differs from the last previously plotted order number by the average number of order numbers required for a halving of value in the three estimates.

Table 1, which is a worksheet of the analysis in the case of an experiment at one Army laboratory, illustrates how plotted points are determined. The worksheet describes calculations to arrive at two chains of points, starting at order numbers 22 and 19, respectively. In the first chain, a halving of value from each of 22, 21, and 23, respectively, involves an average reduction of $6^2/_3$ order numbers. The next plotted point is therefore 22 - $6^2/_3$, or $15^1/_3$; and the bases for the next calculation are 15, 16, and 14. The last point that can be plotted in this first chain is $4^1/_3$.

TABLE 1

Worksheet for Display of Value as a Function of Order Number at an Army Laboratory

Three integers nearest to previous base number	Pairs of order numbers with aggregate value equal to that of first column		Estimates of order number reductions for half values		Base order numbers (plotted points)
	Lower	Higher	Individual	Mean	
			First Chain		
					22
21	13	15	7		
22	15	17	6	$6^2/_3$	$15^1/_3$
23	15	17	7		
14	8	10	5		
15	9	11	5	$5^1/_3$	10
16	9	11	6		
9	2	4	6		
10	4	6	5	$5^2/_3$	$4^1/_3$
11	4	6	6		
			Second Chain		
					19
18	12	14	5		
19	12	14	6	$5^2/_3$	$13^1/_3$
20	13	15	6		
12	6	8	5		
13	6	8	6	$5^1/_3$	8
14	8	10	5		

Figure 2 shows the points of both the first and second chains plotted. Straight lines between adjacent points on each chain provide estimates of value for all order numbers in relation to that of the first plotted point of the chain. The figure also shows a composite line formed by connecting points derived from the plotted points of the first chain and the plotted points of the second chain after a uniform downward adjustment of the latter. The down-

ward adjustment is equal to the average vertical distance between plotted points of each chain and the line connecting points of the other chain.

The linear interpolations between plotted points yield estimates of the relative value to be associated with each order number between the lowest and highest plotted points. Values for still higher order numbers can then be calculated as sums of values of any two previously calculated values (not necessarily separated by two order numbers). Values for the lowest order numbers cannot be calculated by the procedure, but they are likely to be so low that they can be excluded from the program without any worry about loss of bargains.

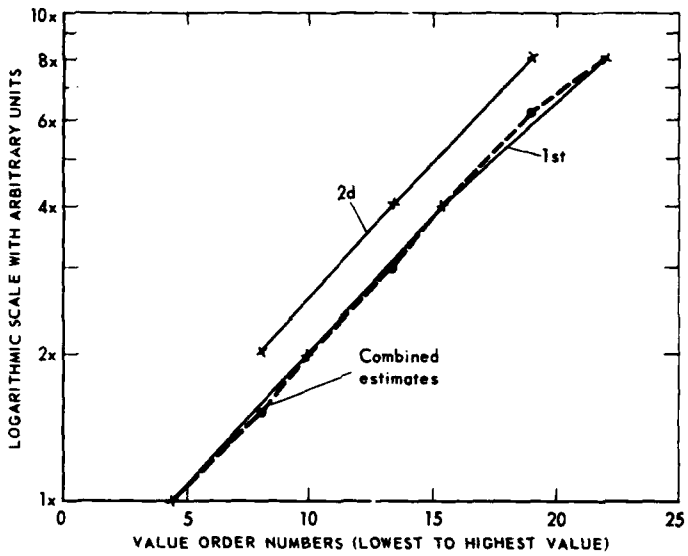

Fig. 2—Value as a Function of Order Number for Proposals at an Army Laboratory

This procedure was tried not only with actual data at Army laboratories, but also at RAC with assumed actual values corresponding to those of a randomized log-normal distribution of values, and the calculated relative values corresponded very closely to the correct relative values. The closeness is indicated by Fig. 3, where the combined estimates are scaled down for comparison with the original data. In the case of the experiment with real project proposals described by the previous two figures, there was no way to check how closely the calculated relative values corresponded to the correct ones. The evaluator who made the ordinal comparisons, however, felt when he was finished that the procedure gave him a set of relative values that he would be willing to use.

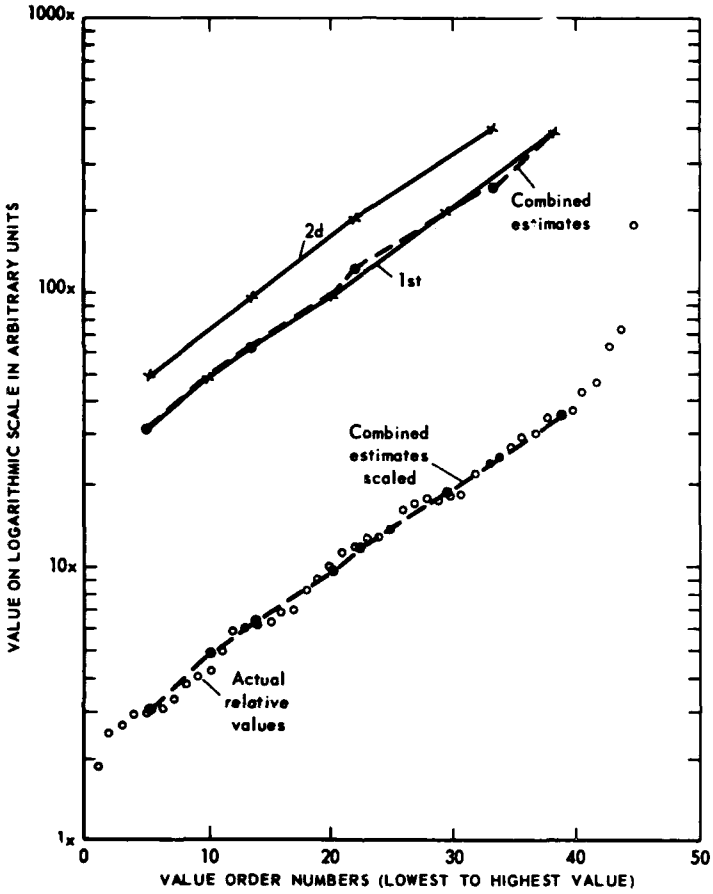

Fig. 3—Value as a Function of Order Number for Hypothetical Data

At another Army installation, an experiment with the same data was made by four evaluators completing their chains independently. A reasonable-looking curve of relation between value and order number was produced by each of the four, but the correlation of rankings among the four was small. They had not been indoctrinated in the structure of value or asked to agree on criteria. As a result, despite the fact that all were high staff members at the same installation, there were serious disagreements on the basic ranking. This points up the fact that, however useful the procedure may be for converting ordinal values to the more useful form of relative values, it must be preceded by good original judgment (aided, perhaps, in some cases by multiple-rating-scale formulas) on the basic ranking of projects.

DEFINITION OF PROJECT PROPOSALS

An unstated assumption underlying all the foregoing discussion of evaluation procedure is that project proposals may be considered discrete, with certain definite (though perhaps unknown) values and costs, and that the values of different proposals can be added to one another.

Those familiar with R&D management know that the scale of effort and the speed of completion of even a well-defined task are variable. It is assumed here, however, that there are bounds to the scales of effort for a budget period that would be considered reasonable and that there are corresponding bounds to the values. It is also assumed that the range of effort and value usually corresponds to a relatively small number of discrete increments in the number of subtasks to be accomplished during the budget period and that, in any case, all the range of possible combinations of value and cost for a project can reasonably be described by a few cases. In such a situation each of the representative cases can be considered a completely separate project for evaluation purposes. The fact that it is not possible to have a program that includes two or more variations of the scope of effort of a single project proposal is a relatively easy problem of the allocation model.

The additivity assumption requires elimination of any substitutability and complementarity of proposals in the list from which selections are to be made. (Substitutability here refers to a situation where the combined value of two or more proposals is less than the sum of values each proposal would have if all the other proposals were excluded from the program. In complementarity, the combined value exceeds the sum of individual values.) Dealing with this problem is also accomplished by proper proposal definition. In general, if two or more proposals have combined values that differ from the sum of the values they would each have in the absence of the others, the list of proposals should be considered to include not only the individual proposals considered independently of one another but also the combinations of proposals that involve significant substitutability or complementarity relations. The allocation model must be designed, of course, to make sure that when related projects are included in a program the values are of the combinations, not sums of individual values.

REFERENCES

REFERENCES CITED

1. Dept of Army, Hq, "Army Research and Development," AR 705-5, 15 Oct 64.
2. Harold Brown, Director of Defense Research and Engineering, Department of Defense, testimony at hearings before House Committee on Appropriations (89th Congress, 1st Session), "Department of Defense Appropriations for 1966," Pt 5, US Govt Printing Office, Washington, D.C.
3. C. G. Suits, "Selectivity and Timing in Research," Res. Mgt., p 418 (Nov 62).
4. Col Lee R. Standifer, "From Concept to Application," Air University Quart. Rev., Winter and Spring, p 76 (1962–1963).
5. E. Berman, Research Analysis Corporation, unpublished data.
6. P. C. Fishburn, "Independence in Utility Theory with Whole Product Sets," Opn. Res., 13: 28–45 (Jan.–Feb. 65).
7. C. W. Churchman and R. L. Ackoff, "An Approximate Measure of Value," Opn. Res., 2: 172–87 (May 54).
8. M. J. Peck and F. M. Scherer, "Technical Notes on the Evaluation Experiment," App 19A of The Weapons Acquisition Process: An Economic Analysis, Harvard University Press, Cambridge, Mass., 1962, pp 669–704.
9. L. I. Thurstone and L. V. Jones, "The Rational Origin for Measuring Subjective Values," J. Am. Stat. Assn., 52: 457–71 (Sep 57).
10. P. C. Fishburn, "Methods of Measuring Additive Utilities under Independence," unpublished notes, Research Analysis Corporation, 1965.

ADDITIONAL REFERENCES

Baker, N. R., and W. H. Pound, "Research Project Selection: Where We Stand," IEEE Trans. Engr. Mgt. (1965).

——, Department of Industrial Engineering and Management Sciences, The Technology Institute, Northwestern University, Evanston, Ill., mimeograph.

Dean, Burton V. (ed.), Opn. Res. in Res. and Dev., Proceedings of a Conference, Case Institute of Technology, John Wiley & Sons, Inc., New York, 1963.

Hitch, C. J., and R. N. McKean, The Economics of Defense in the Nuclear Age, Harvard University Press, Cambridge, Mass., 1960.

National Bureau of Economic Research, "The Rate and Direction of Inventive Activity: Economic and Social Factors," a Conference of the Universities' National Bureau Committee for Economic Research and the Committee on Economic Growth of the Social Science Research Council, Princeton University Press, Princeton, N. J., 1962.