AD-A047 479 HONEYWELL INC MINNEAPOLIS MINN SYSTEMS AND RESEARCH --ETC F/G 9/2
DISTRIBUTED DATA PROCESSING TECHNOLOGY. VOLUME VIII. APPLICATIO--ETC(U)
SEP 77 R PENNINGTON, K WATSON, G CONVOY DASG60-76-C-0087
UNCLASSIFIED 77SRC72 NL

| OF |
AD
A047479

END
DATE
FILMED
1 -78
DDC

2

DISTRIBUTED DATA PROCESSING TECHNOLOGY

DASG60-76-C-0087

FINAL REPORT

VOLUME VIII

APPLICATION OF DDP TECHNOLOGY TO BMD: RESEARCH PERFORMANCE MEASUREMENT

For
BMDATC
Ballistic Missile Defense
Advanced Technology Center
Huntsville, Alabama 35807

Honeywell
SYSTEMS & RESEARCH CENTER
2600 RIDGWAY PARKWAY
MINNEAPOLIS, MINNESOTA 55413

UNCLASSIFIED

SECURITY CLASSIFICATION OF THIS PAGE (WHEN DATA ENTERED)

REPORT DOCUMENTATION PAGE

READ INSTRUCTIONS BEFORE COMPLETING FORM

1. REPORT NUMBER

2. GOV'T ACCESSION NUMBER

3. RECIPIENT'S CATALOG NUMBER

9

4. TITLE (AND SUBTITLE)

6

Distributed Data Processing Technology.
Volume VIII. Application of DDP Technology
to BMD: Research Performance Measurement.

5. ~~TYPE OF REPORT/PERI~~OD COVERED

Final Report.
Octo[redacted] 1976 ~~to~~ Oct[redacted] 1977,

6. ~~PERFORMING ORG. REPORT NUMBER~~

14

77SRC72

7. AUTHOR(S)

10

R./Pennington, K./Watson, G./Convoy
B./Stewart

8. CONTRACT OR GRANT NUMBER(S)

15

DASG60-76-C-0087

9. PERFORMING ORGANIZATIONS NAME/ADDRESS

Honeywell Systems & Research Center,
Minneapolis, MN/General Research Corporation,
Santa Barbara, CA

10. PROGRAM ELEMENT, PROJECT, TASK AREA & WORK UNIT NUMBERS

11. CONTROLLING OFFICE NAME/ADDRESS

Ballistic Missile Defense
Advanced Technology Center
Huntsville, AL 35807

12. REPORT DATE

11

Sep[redacted]77

13. NUMBER OF PAGES

62

14. MONITORING AGENCY NAME/ADDRESS (IF DIFFERENT FROM CONT. OFF.)

12 62p.

15. SECURITY CLASSIFICATION (OF THIS REPORT)

Unclassified

15a. DECLASSIFICATION DOWNGRADING SCHEDULE

16. DISTRIBUTION STATEMENT (OF THIS REPORT)

Approved for public release; distribution unlimited.

DDC
RECEIVED
DEC 8 1977
F

17. DISTRIBUTION STATEMENT (OF THE ABSTRACT ENTERED IN BLOCK 20, IF DIFFERENT FR

18. SUPPLEMENTARY NOTES

19. KEY WORDS (CONTINUE ON REVERSE SIDE IF NECESSARY AND IDENTIFY BY BLOCK NUMBER)

Technology assessment
Research performance measurement
Distributed data processing
Design technology evaluation models
Comparative research evaluation
Goal achievement research evaluation

20. ABSTRACT (CONTINUE ON REVERSE SIDE IF NECESSARY AND IDENTIFY BY BLOCK NUMBER)

This report is concerned with measurement and evaluation of Distributed Data Processing (DDP) research and the products of that research. A measure of completeness and conformity to the statement of work is included. Two design technology evaluation models are presented. An evaluation of the literature on research measurement is discussed, and a particular set of criteria and a research measurement approach are suggested.

HD-168 REV 11/74

DD FORM 1 JAN 73 1473 EDITION OF 1 NOV 55 IS OBSOLETE

UNCLASSIFIED

SECURITY CLASSIFICATION OF THIS PAGE (WHEN DATA ENTERED)

402 349

LB

SECURITY CLASSIFICATION OF THIS PAGE (WHEN DATA ENTERED)

SECURITY CLASSIFICATION OF THIS PAGE (WHEN DATA ENTERED)

FOREWORD

The research documented in this volume was conducted under Ballistic Missile Defense Advanced Technology Center contract number DASG60-76-C-0087, entitled "Distributed Data Processing Technology." The work was performed by Honeywell Systems and Research Center, Minneapolis, Minnesota, and General Research Corporation (GRC), Santa Barbara, California under the direction of Mr. C. R. Vick, Director, Data Processing Directorate, Ballistic Missile Defense Advanced Technology Center. Mr. J. Scalf was the BMDATC project engineer for this contract; Ms. B. C. Stewart was the Honeywell/GRC program manager.

This report covers work from October 1976 to October 1977. It represents a Distributed Data Processing (DDP) Performance Measurement (PM) Plan that defines means to measure the degree of success for each of the distinct research areas presented in other volumes of this report. These PMs include: 1) the measure of improved performance achieved with a given solution as compared with the projected improvement; 2) the identification of research requirements derived from the research objectives (which provide measures of success for the development of design concepts such as DDP architectural design methods and procedures, simulation languages and design automation tools); and 3) models which provide estimates of desired performance.

Midwest Research Institute contributed to this volume on behalf of Honeywell. R. Pennington of GRC and B. C. Stewart of Honeywell wrote portions of this report, and the final version was written by B. C. Stewart.

This document is Volume VIII of the final report. Other volumes of the report are the following:

Volume I - Management Summary

Volume II - DDP Rationale: The Program Planning Point of View

Volume III - DDP Rationale: The Technology Point of View

Volume IV - Application of DDP Technology to BMD: Architectures and Algorithms

Volume V - Application of DDP Technology to BMD: DDP Processing Subsystem Design Requirements

Volume VI - Application of DDP Technology to BMD: Impact on Current DP Subsystem Design and Development Technologies

Volume VII - Application of DDP Technology to BMD: Experiments

Volume IX - DDP Rationale: The Program Experience Point of View

---

*Volumes V, VI, the appendix to Volume VII, and one section of Volume VIII were prepared by General Research Corporation.

CONTENTS

CONTENTS (continued)

LIST OF TABLES

SECTION 1
INTRODUCTION

As specified in the Foreword, the purpose of this volume is to provide performance measures established to measure progress and to provide data from studies, tests, and demonstrations which quantify actual progress.

As prescribed in Paragraph 3.9 of SW-A-131-76, performance measures of the following types are to be considered:

- The measure of improved performance achieved with a given solution compared with the projected improvement.
- The identification of the research requirements derived from research objectives which provide measures of success for the development of design concepts such as Distributed Data Processing (DDP) architectural design methods and procedures, simulation languages, and design automation tools.
- Models which provide estimates of desired performance.

It is important to note that in the above statements the term "performance" is used with relation to three entirely different objects. These objects are:

- The research program itself
- The development methodology for distributed DP systems for Ballistic Missile Defense (BMD)
- The resulting BMD system

The basic objective is to derive performance measures for the research program itself. How well has the research progressed? What has it accomplished? Is the BMD potential for the United States enhanced by the performance of this research? If so, in what ways and how much? We will refer to the measures that directly relate to these sorts of questions as "research performance measures."

To address the above questions it is necessary to consider other levels of questions. These relate to the objectives of the research. What do we expect the research to accomplish? These expectations fall into two classes. The first class relates to the design and development procedures for BMD systems. The research may enable us to do a better job in developing a BMD system--develop it more quickly, more cheaply, on a more predictable cost, and with more predictable results. To decide whether the research has in fact contributed in this area, we must address the "performance" of the development process itself. How well can we develop a BMD system now? How well do we expect to be able to develop a BMD system if the research is successful?

The second class of expectations for the research has to do with the attributes of the BMD Distributed Processing (DP) subsystem itself. The research may enable us to develop a DP subsystem having superior attributes. The system may enable the BMD system to counter more-sophisticated or more-massive threats. It may be more survivable, more reliable, more easily operated and/or maintained, more capable of incremental growth or improvement to meet expanding or new requirements. To decide whether the research has in fact contributed in this area, we must address the "performance" of potential BMD DP subsystems produced by current methods versus those produced by new methods growing out of this research. How good a BMD DP subsystem can we develop now? How good a BMD DP subsystem will we be able to develop if the continuing research is successful?

From the previous discussion it can be seen that evaluating the success of the research program requires the assessment of performance for:

- Current and projected BMD development methodologies
- BMD DP subsystems resulting from current and projected methodologies

It is immediately clear that complete quantitative performance assessments along the above lines would be prohibitively expensive, if not impossible. However, the basic approach is sound and enforceable. As is customary in cases where empirical data are not available, the BMD development methodologies and resulting DP subsystems will have to be represented by hypothetical models. These models will be constructed so as to contain as quantified attributes those features that have in the past been qualitatively adjudged to be most amenable to improvement by the application of DDP techniques. Wherever possible, the models can be calibrated against experience data.

Because of the wide range and differences among the research tasks documented in this final report, several different approaches to research performance measurement have been pursued and are presented in this Volume.

Section 2 is concerned with the consistency and completeness of the overall research performed with respect to the requirements of the Statement of Work.

Section 3 presents one modelling approach for performance measurement of a DDP Design Methodology.

Section 4 summarizes and evaluates the literature and existing approaches to research performance measurement.

Section 5 discusses the applicability of existing measurement approaches to research performed on DDP and makes recommendations concerning approaches which best suit the research documented in other volumes of this report.

SECTION 2

MEASUREMENT OF COMPLETENESS OF RESEARCH PERFORMED

The Performance Measure (PM) for the overall completeness of the research is a cross-reference table (Table 1) which summarizes the research requirements from the DDP Statement of Work and points to the section of the final report which addresses those requirements.

TABLE 1. COMPLETENESS MEASURE CONFORMANCE MATRIX

Research Requirement from Statement of Work	Volume/Page Reference
I. Distributed Computer Architecture Studies:	
1) Conventional Architecture:	
a) Identify conventional architectures.	III, IX
b) Identify BMD applications for conventional architectures.	III, IV
c) Identify deficiencies for architecture/application pair.	III, IV
2) Advanced Configurations:	
a) Define solutions to reduce deficiency.	IV
b) Identify application for new architectures.	IV
c) Define experiments to quantify success.	VII
II. DDP Algorithm Architecture Studies:	
a) Identify and define algorithms.	IV
b) Determine architecture for which it is applicable.	IV
c) Recommend real-time implementation	IV
III. DDP Subsystem Architecture Requirements Studies:	
1) Node Characteristics:	
a) Establish partitioning criteria for assigning BMD operations to nodes.	V, II,
b) Define rules for both static and dynamic control.	V
c) Define methods of fault detection.	V, IV
2) Networks:	
a) Define node structure.	V
b) Define link structure to connect nodes with DDP subsystem architectures.	V
c) Define approaches to ensure network integrity.	V
3) Data Base:	
a) Design criteria for partitioning/assigning data to nodes.	V, VII, II
b) Define concepts to ensure data integrity.	V, VII
4) Performance:	
a) Define performance issues for network of nodes.	V, VI, II
b) Define the impact of performance validation and SETS.	V, VI

TABLE 1. COMPLETENESS MEASURE CONFORMANCE MATRIX (continued)

Research Requirement from Statement of Work	Volume/Page Reference
IV. Distributed Node Architectural Requirements Studies:	
1) Process Characteristics:	
a) Establish partitioning criteria for assigning operations to processes.	IV, V
b) Define rules for both static and dynamic control.	IV, V
c) Define ways of ensuring process integrity.	IV, V
2) Networks:	
a) Define process structure.	IV, V
b) Define link structure.	IV, V
c) Define approaches to ensure network integrity.	IV, V
3) Data Base:	
a) Design criteria for partitioning/assigning data to processes.	IV, V
b) Define concepts to ensure data integrity.	IV, V
4) Performance:	
a) Define performance issues for network of processes.	V, IX
b) Define impact of performance validation and SETS.	VI
V. Real-Time Engagement Software/Hardware Architectural Requirements Studies:	
1) Hardware Architecture:	
a) Define hardware attributes for DDP architectures.	III, IV, IX
2) Real-Time Engagement Software Architecture:	
a) Define impact of selecting real-time engagement software.	VI
VI. Design and Development Technology:	
1) Software Engineering Technology:	
a) Quantify critical deficiencies of current software engineering technology.	VI
b) Identify approaches to solving deficiencies.	VI
c) Establish plan for developing required technology.	VI
2) Performance Validation Technology:	
a) Identify performance measures for evaluating and validating DDP subsystems.	VI
b) Identify and quantify critical deficiencies.	VI
c) Analyze approaches to solving deficiencies.	VI
d) Develop research plan for resolution.	VI
3) SETS Technology:	
a) Determine the architecture of SETS for DDP application.	VI
b) Identify tradeoffs of hardware/software requirements.	VI
c) Identify critical related technology issues.	VI
d) Develop research plan.	VI

TABLE 1. COMPLETENESS MEASURE CONFORMANCE MATRIX (concluded)

Research Requirement from Statement of Work	Volume/Page Reference
VII. DDP Experiments:	
1) General Requirements	VII
a) Engineering description	VII
b) Issues being addressed	VII
c) Objectives	VII
d) Schedule of performance	VII
e) Performance measures for evaluating results	IX
2) Experimental Software Requirements:	
a) Purpose of package	VI, VII
b) Description of environment	VI, VII
c) Description of functions	VI, VII
d) Relationship to other software	VI, VII
e) Data requirements	VI, VII
f) Implementation constraints	VI, VII
g) Definition of functional performance capabilities	VI, VII
3) Experimental Hardware Requirements:	
a) Definition of number, type, and organization of equipment	VII
b) Description of network configuration	VII
c) Description of functions of equipment	VII
d) Description of interfaces	VII
e) Definition of functional performance capability	VII
VIII. Performance Measures for DDP Research:	
1) Measures of Improved Performance	VIII
2) Measures for the Success of Design Concepts:	
a) Identification of research objectives	VIII
b) Derivation of research requirements	VIII
c) Development of measures to evaluate achievement of objectives	VIII
3) Models Which Provide Estimates of Desired Performance	VIII
IX. DDP Payoff:	
1) Establish Payoff of DDP With Respect to Centralized Processing	II, IX

SECTION 3

MODELS FOR EVALUATION OF DDP DESIGN METHODOLOGY

3.1 DDP DESIGN METHODOLOGY MODELING

A number of techniques are in existence for the quantitative modelling of the cost and schedule features of a development program. The best-known of these, Program Evaluation and Review Technique (PERT) and Critical Path Management (CPM), were developed as management tools for the control of large-scale, high-technology development programs. This is a purpose considerably different from ours. We wish only to describe, rather than control, such programs. Nevertheless, the quantitative features of interest are the same for both intents. We are interested in the amount of time and the cost to develop a DP subsystem following one methodology versus that required to develop a DP subsystem meeting the same requirements following a different methodology. (Note that we have not spoken of two methods for developing the same system, but rather of two methods for developing systems from the same set of requirements. We start from one set of requirements, but we expect different development methods to result in different systems.) The same types of activity networks and event networks used for CPM and PERT will also suffice for our purposes. Therefore, we will base our modelling of the DDP design methodology on these techniques, as described in Reference 1.

The first step in developing the activity network for a project is to develop an activity table. The activity table contains a line entry for each activity in the development program. The line entry contains an activity name, description, list of immediate predecessor activities, time duration, and cost (in more-sophisticated models the latter two entries may be replaced by a tradeoff curve). Using standard techniques, the table, when completed, may be converted to Activity-on-Node (AON) charts or arrow diagrams. In turn, total time requirements and costs can be determined, together with estimates of variances.

For a starting point on the current program, two development methodology models are required. The first is a "control" model, representing the way in which we would expect a distributed system would be designed if no research were done to improve the design methodology. We will base this model on the BMDATC Software Development System, as the latest available documented system. The second is a "goal" model, representing the way in which we hope to be able to develop a system if the research is completely successful. We will base this on the baseline design method described in Volume 5 of this report. As the Phase 2 DDP research progresses, additional models will be developed. These new models will fall into two classes. The first class will be revised "goal" models, representing the baseline design method as it is revised to account for things learned during the research. The second class will be revised "current state-of-the-art" models, representing the SDS method as adjusted to incorporate features that have been shown by ongoing studies, tests, or demonstrations to be well-enough established for inclusion in the baseline method.

From these four model classes, several figures of merit can be computed. Let

T_c = total development time using the control model

T_g = total development time using the goal model

T_r = total development time using the current version of the revised goal model

T_n = total development time using the current version of the current state-of-the-art model

Then we have the following research performance measures:

T_g/T_c = conceptual development time compression achievable through the research

T_n/T_c = development time compression already achieved by the research

T_r/T_n = potential development time compression still achievable through continuing the research

In like manner, let

C_c = total development cost using the control model

C_g = total development cost using the goal model

C_r = total development cost using the current version of the revised goal model

C_n = total development cost using the current version of the current state-of-the-art model

Then we have the following research performance measures:

C_g/C_c = conceptual development cost reduction factor achievable through the research

C_n/C_c = development cost reduction factor already achieved by the research

C_r/C_n = potential development cost reduction factor still achievable through continuing the research

In like manner, let

ΔT_c = uncertainty in total development time using the control model

ΔT_g = uncertainty in total development time using the goal model

ΔT_r = uncertainty in total development time using the current version of the revised goal model

ΔT_n = uncertainty in total development time using the current version of the current state-of-the-art model

Then we have the following research performance measures:

$100\left(1 - \frac{\Delta T_g T_c}{\Delta T_c T_g}\right)$ = percentage improvement in development time uncertainty conceptually achievable through the research

$100\left(1 - \frac{\Delta T_n T_c}{\Delta T_c T_n}\right)$ = percentage improvement in development time uncertainty already achieved by the research

$100\left(1 - \frac{\Delta T_r T_n}{\Delta T_c T_n}\right)$ = percentage improvement in development time uncertainty still potentially achievable through continuing the research

In like manner, let

ΔC_c = uncertainty in total development cost using the control model

ΔC_g = uncertainty in total development cost using the goal model

ΔC_r = uncertainty in total development cost using the current version of the revised goal model

ΔC_n = uncertainty in total development cost using the current version of the current state-of-the-art model

Then we have the following research performance measures:

$100\left(1 - \frac{\Delta C_g C_c}{\Delta C_c C_g}\right)$ = percentage improvement in development cost uncertainty conceptually achievable through the research

$100\left(1 - \frac{\Delta C_n C_c}{\Delta C_c C_n}\right)$ = percentage improvement in development cost uncertainty already achieved by the research

$100\left(1 - \frac{\Delta C_r C_n}{\Delta C_c C_n}\right)$ = percentage improvement in development cost uncertainty still potentially achievable through continuing the research

In constructing the activity tables for the development models, we must have as a starting point some set of initial requirements statements for the system. The requirements statements must give a sufficiently complete description to allow quantitative values to be entered in the cost columns in the activities tables.* For this purpose we have chosen to use the set given in Reference 2. This choice is motivated by the fact that it is the only one immediately available of sufficient completeness that is not oriented toward a single site system.

The activity table for the control model, SDS, is given in Table 2. The activities listed are extracted directly from the TRW Software Requirements Engineering and the TI Process Design Engineering Documentation produced for SDS. The time and cost estimates for the individual activities are primarily judgment values, tempered by numerical factors extracted from cost-estimating relationships developed by GRC under contract to NASA. Based on this activity table, the estimated development times and costs and their uncertainties are:

T_c = total development time using the control model

C_c = total development cost using the control model

ΔT_c = uncertainty in total development time using the control model

ΔC_c = uncertainty in total development cost using the control model

---

[*] This is a requirement on the output of the Axiomatic Requirements Engineering Research program.

TABLE 2. RESEARCH PRODUCT MODEL TO REPRESENT DDP DEVELOPMENT--EXAMPLE ACTIVITY TABLE FOR SDS

Job Identification	Job Description	Immediate Predecessors	Time to Perform (mos)	Cost to Perform ($M)
A	Define DPS elements	---	3	4
B	Evaluate results with REVS-RADX	A	2	6
C	Complete functional definition (define R-Nets)	B	12	10
D	Develop/use functional models	C	10	10
E	Add management segment	C	6	12
F	Identify performance requirements	---	6	8
G	Locate test points on R-Nets	D, E, F	3	8
H	Define initial validation paths	G	3	2
I	Complete validation path definition	H	12	6
J	Develop analytic models	D, E	12	20
K	Regroup requirements into tasks	I	6	8
L	Assign tasks to processors	D, I	3	6
M	Sequencing logic design	L	6	8
N	Structure real-time algorithms	M, J	12	40
O	Process design validation	M	6	12

The activity table for the goal model (Table 3) is extracted from the preliminary development plan, CDRL Item A005. Again, the time and cost estimates for the individual activities are primarily judgment values. Based on this activity table, the estimated development times and their uncertainties are:

T_g = total development time using the goal model
C_g = total development cost using the goal model
ΔT_g = uncertainty in total development time using the goal model
ΔC_g = uncertainty in total development cost using the goal model

From these estimates, initial estimates can be given for some of the research performance measures, as follows:

- Conceptual development time compression achievable through the research =
- Conceptual development cost reduction factor achievable through the research =
- Percentage improvement in development time uncertainty conceptually achievable through the research =
- Percentage improvement in development cost uncertainty conceptually achievable through the research =

Numerical values for others of the research performance measures related to development will not be computable until some of the research tasks have actually been completed.

TABLE 3. EXAMPLE PERFORMANCE MEASURE VALUES FOR DDP DEVELOPMENT

A. Development Attributes

Parameter	Control Model	Goal Model	Current State-of-the-Art Model	Revised Goal Model
Development time	60 mos	36 mos	54 mos	40 mos
Development cost	$150M	$75M	$140M	$100M
Development time uncertainty	24 mos	12 mos	22 mos	18 mos
Development cost uncertainty	$60M	$25M	$50M	$40M

B. Research Performance Measures

Parameter	Conceptually Achievable Through Research	Already Achieved by Research to Date	Still Potentially Achievable by Continuing Research
Development time compression	0.6	0.9	0.74
Development cost reduction	0.5	0.93	0.71
Improvement in development time uncertainty	18%	-2%	10%
Improvement in development cost uncertainty	17%	12%	-12%

3.2 MODELING OF THE RESULTING SYSTEM ATTRIBUTES

In past years, BMD systems in wide variety have been represented by models ranging from very high-level strategic exchange models to extremely detailed emulations of the BMD computer hardware and software. We are not aware, however, of any model having the specific properties required for our current purposes.* The model we seek must provide for quantitative evaluation of the following system attributes:

- BMD effectiveness
- Reliability
- Survivability
- Maintainability
- Availability
- Cost
- Growth capability
- Deployability (speed with which the system can be deployed)
- Robustness (degree to which the system can continue to function in a degraded mode after loss of some system elements)
- Predictability (the degree to which the system will actually perform as it was intended to perform)

For most of these attributes, a distributed system offers immediate and obvious advantages vis-a-vis a centralized system. The BMD payoffs

---

*This is a requirement for the Axiomatic Requirements Engineering Research program.

for distributed processing are treated separately under Task 3 of the contract, and the results are reported in CDRL Item A004. What we seek to quantify here seems to be at another level of refinement--the efficacy of a distributed system developed under one methodology vis-a-vis that of a distributed system developed under another methodology. In one sense this seems a hopeless task. We need a system model of a type not previously developed, to model systems of a type never yet constructed. In another sense, the task is somewhat approachable because distributed processing for BMD systems has not yet been performed. This fact alone tells us that a first distributed system developed under the SDS approach is likely to rate low on effectiveness, cost, maintainability, growth capability, deployability, robustness, and predictability, merely because the SDS procedure has not been applied to a distributed system with quantified goals for these attributes. Thus, we are justified in predicting a positive return for the research itself, regardless of the degree of credibility we can muster for the new type model for new type systems that will be presented below.

The approach taken in developing the model structure is to insist that for each of the system attributes listed earlier, the model must be capable of producing a numerical value. This can be done by defining a set of structural elements for the model, defining a method for interrelating these structural elements to represent a specific system, and defining the formulas to be used with respect to these elements to generate numerical values for the system attributes. There is a potential nomenclature problem that we wish to avoid in setting up this structure. The research itself is concerned with giving precise meanings to such terms as nodes, processes, and resources. Our model must treat entities that are in some sense related to entities in the real or virtual systems being treated under the research. To avoid unintended confusion of terms, or unintended biasing of terms used for describing actual systems, we will studiously avoid using terms such as node, process, or resource in connection with the models. We will deal with entity types designated only by a type designator and a set of properties.

The fundamental working elements of the system will be designated as Type A entities. The Type A entities for a given system fall into some prescribed number of subtypes, A1, A2, A3, etc. Each subtype is characterized by a set of parameter values. For A_i these are:

a_i = cost value for the entity
n_i = interconnect value for the entity
f_i = flow value for the entity
x_i = complexity value for the entity
u_{ij} = consumptive value for the entity with relation to entity B_j (defined below)

Clearly, we intend Type A entities to represent some quantities of MBD software, but we refuse at this point to be any more specific about their nature.

The next set of elements of the system will be designated as Type B entities. These also fall into some prescribed number of subtypes, B1, B2, B3, etc. Each subtype is characterized by a set of parameter values. For B_i these are:

b_i = cost value for the entity
v_i = capacity value for the entity
r_i = reliability value for the entity
s_i = availability value for the entity

Individual Type A entities can be assigned to Type B entities, so long as

$$\sum_i uij \leq v_j$$

Clearly, the Type B entities are intended to represent hardware processing units of some sort, but, again, we refuse at this point to be more specific about the nature of the representation.

The next set of elements will be designated as Type C entities, again divided into Subtypes C1, C2, etc. The parameters for C_i are:

c_i = local cost value for the entity
z_i = distance cost value for the entity
w_i = flow rate value for the entity
t_i = delay time for the entity

Type C entities will in some way represent hardware interconnection paths. A Type C entity can be assigned to a pair of Type B entities.

The next set of elements will be designated as Type D entities, again divided into Subtypes D1, D2, etc. The parameters for D_i are:

d_i = cost value for the entity
s_i = survivability value for the entity

Type D entities will somehow represent sites located at sufficient distance from each other to be independent from the vulnerability standpoint. Individual Type A entities can be assigned to Type D entities without limit.

A complete system is defined by giving the parameter values for all of the allowed subtypes of Types A, B, C, and D, then prescribing and identifying the individual members, or instances of each subtype. Then, each instance of a B subtype must be assigned to one instance of a D subtype, and each instance of an A subtype must be assigned to at least one instance of a B subtype. Other assignments are made as appropriate to give the system the intended structure. Once this is done, the performance measures for the system can be computed in accordance with the procedures given below.

3.2.1 Effectiveness

In a real BMD system, the effectiveness is related to the ability to find, track, and destroy threatening objects. Most models treat this portion of the system operation with an attempt at realism, with higher fidelity usually considered better. For our purposes, we choose to address effectiveness on a much more abstract basis, and have only a heuristic relation to "realistic" effectiveness measures. For this purpose, we note that the operations of search, track, and kill generally invoke sequences of software routines, with the requirement that the sequences be executed within limited time periods to keep up with real-world events. We note also that the number of objects that can be killed relates to the number of software sequences the system can execute simultaneously. With this in mind, we define system effectiveness for our model system so as to have properties analogous to the ones just described.

System effectiveness will be assumed to be related to the capability to chain rapidly through sequences of Type A entities. For simplicity in this initial application, it will be assumed that there is only one chain of interest, say A1-A2-A3-A4-A5. If the approach proves successful, supplemental chains can be added. A sequence is chained through by the following steps:

1) Find an instance of A1 in the system.
2) Reduce the value of v_j at the B_j on which this A1 resides by an amount of $u1_j$.
3) Compute a time $\tau = u1_j / v_j^o$, where v_j^o is the original value of v_j

4) Find an instance of A2 in the system.

5) If A2 is at a different B_k:

    a) Find a path of C_n's from B_j to B_k;

    b) On each of the C_n's, reduce the value of w_n by an amount f1.

    c) To τ add the T_n's for each C entity on the path.

6) Reduce the value of v_k at the B_k on which this A2 resides by an amount of $u2_k$.

7) Increase τ to $\tau = \tau + u2_k / v_k^o$.

8) Find an instance of A3 in the system, etc.

This is continued until a complete chain A1-A2-A3-A4-A5 has been found. At this point we have a value of τ for this chain and the system network modified by the reduction of some of the v_j's and w_j's. An attempt is made to repeat this process until no additional chains can be found without making some v_j or w_j negative. The effectiveness is defined to be the maximum number of chains that can be created in this manner such that the τ for each chain is less than some prescribed value.

3.2.2 Reliability

For a real system, reliability is usually taken as the probability that, given that the system is performing satisfactorily at the beginning of a mission, it will perform satisfactorily throughout the mission. The term "satisfactorily" requires definition for the particular system involved. The only failure cause considered is random component failure. In our system, success of operation is equated to the existence of the chains of Type A entities as described in the preceding section. Thus, for our purposes, reliability is defined to be the probability that, given the reliabilities r_i

of the individual B_i's, the number of chains remaining operable for a 30-minute period is at least 90 percent of the number found for the full system.

3.2.3 Survivability

Survivability is a measure of the capability of a system to resist the depreciative effects of an enemy attack. This in turn should relate to the number of individual aim points, or sites, that have to be removed to disable the system. In our system, the threatened units for survivability purposes are the Type D entities. Thus, we find it reasonable to define survivability as the number of Type D entities that can be randomly removed from the net before the effectiveness drops to 50 percent of its normal value.

3.2.4 Maintainability

Maintainability of a system depends on many factors. Among these are:

- The number of different system elements involved
- The complexity of the individual elements
- The complexity of the interconnections among elements
- The degree of saturation under which the system must operate

We seek a heuristic formula that exhibits the appropriate behavior pattern with respect to each of these. The x_i parameters for the A_i subtypes are presumably the measures of complexity for individual software components. It is well known that software becomes more expensive as it approaches saturation of the equipment on which it operates. Thus, the x_i's should be corrected by a factor that increases as the equipment approaches saturation. Considerations such as these have led to the formula described below.

Maintainability is defined to be obtained by the following steps: Consider the chain-forming operation of Section 3.2.1. Let v_i be the capacity of B_i before any chains are formed; and v_{im} be the capacity of B_i after m chains are formed; let w_i be the flow rate of C_i before any chains are formed; and w_{im} be the flow rate of C_i after m chains are formed.

1) For each of the chains found in the effectiveness calculation, form the sum:

$$\xi_m = \sum_i x_i \left(e^{1 - \frac{v_{im}}{v_i}} + \delta_i^{i+1}\, \mathrm{nie}^{1 - \frac{w_{im}}{w_i}} \right)$$

where

$$\delta_i^j = \begin{cases} 0 & \text{if } A_i \text{ and } A_j \text{ are assigned to the same B} \\ 1 & \text{otherwise} \end{cases}$$

2) Form the quantity:

$$x = \left(\sum_m \xi_m e^{\tau_m/T-1} \right) (yq)^{1/2}$$

where

τ_m = time requirement for chain m

T = allowed time for a chain

y = number of different B subtypes in the system

g = number of subtype C instances divided by the number of subtype B instances in the system.

Then the maintainability measure is defined as M = 1/x.

3.2.5 Availability

Availability for a real system is usually taken to be the probability, given the reliabilities and mean times to repair of individual subsystems, that the system will be operable and ready to perform its mission at any arbitrarily selected moment. Having assigned availability values to our individual Type B entities, we can model availability with a formulation similar to that used for reliability.

Availability is defined to be the fraction of the time, considering the availability of the individual B_i's, that the number of chains available is at least 90 percent of the number found in the efficiency computation for the whole system.

3.2.6 Cost

Except for software, the system cost is taken to be the sum of the individual costs of the elements in the system. Because of the strong dependency of software costs on the integration aspects, and particularly on the degree to which the hardware resources are saturated, we wish to correct the software costs for this sort of effect. Since maintainability is affected by the same sorts of factors, we do this by multiplying the software cost term by the reciprocal of the maintainability measure. Thus, we have the cost defined as:

$$\text{Cost} = \left(x \sum_{i} a_i + \sum_{j} \beta_j b_j + \sum_{k} \gamma_k c_k + \sum_{\ell} \delta_\ell d_\ell + \sum_{k} \epsilon_m^n z_k \right) / E$$

where

x = reciprocal of maintainability measure (see 3.2.4)

β_j = number of instances of type B_j entities in the system

γ_k = number of instances of type C_k entities in the system

δ_ℓ = number of instances of type D_ℓ entities in the system

$$\epsilon^n_m = \begin{cases} 0 \text{ if the pair of entities of Types } B_j \text{ and } B_k \text{ to which } C_k \\ \text{if assigned belong to the same instance of D.} \\ 1 \text{ otherwise.} \end{cases}$$

E = effectiveness as defined in 3.2.1.

Thus, our "cost" is a normalized quantity, normalized to unit effectiveness.

3.2.7 Growth Capability

Growth capability is another system attribute that is difficult to quantify. Clearly, growth will be easiest if system elements are small, uniform in construction, and independent of each other in their operation, so that the addition of new elements could be done in increments of any size without disturbing already deployed portions of the system. Growth might be accomplished by adding new sites or by adding capability to existing sites. Our initial cut at an index will consider only the addition of new sites.

For the treatment of growth by adding new sites, perform the following steps:

1) Increase the total values of

$$\sum_i \beta_i v_i \text{ and } \sum_i \gamma_i w_i$$

in the system be 20 percent by replicating instances of D_i subtypes.

2) Compute the effectiveness with the constraining assumption that all chains from the old system must be retained. Call this E1.

3) Recompute the effectiveness without constraining the choice of chains. Call this E2. Define the growth capability measure G to be

$$G = \frac{E2 - 1.2E}{1.2E} - \frac{E2 - E1}{1.2E}$$

3.2.8 Deployability

The deployability attribute carries many of the same sorts of features as the growth capability attribute. The main difference is that in considering deployability we must consider being able to reach some minimum capability from a zero-capability stance. This will be enhanced if the system elements can be built and tested in small units and then put on the shelf. We will approximate this by computing an average unit cost associated with creating one of the processing chains constructed in the effectiveness calculation, this cost being associated with the cost of processors, sites, and interconnections involved in providing the environment for the chain.

Deployability is defined to be obtained by the following steps:

1) For each chain in the chain set derived in the effectiveness calculation, compute a deployment cost increment as follows:

    a) For each of the A_i in the chain, compute its environment cost as:

$$\epsilon_i = b_j + d_k + \delta_\ell^m z_\ell$$

where

$$\delta_\ell^m = \begin{cases} 0 \text{ if } A_i \text{ and } A_{i+1} \text{ belong to the same D.} \\ 1 \text{ otherwise.} \end{cases}$$

b_j = the B type entity to which A_i belongs.

d_k = the D type entity to which b_j belongs.

b) Compute a chain environment cost as:

$$\eta = \Big(\sum_i \xi_i\Big)(\text{number of distinct B-type instances represented})^{1/2} \times (\text{number of distinct D-type instances represented})^{1/2}$$

2) Find the average of the chain environment costs for all chains. Call this the deployment cost increment.

Deployability is then defined to be the reciprocal of the deployment cost increment.

3.2.9 Robustness

Robustness is a measure of the capability of the system to accept punishment and still function successfully. Clearly, this relates in some degree to survivability (see 3.2.3), but with a somewhat different twist. We wish to know how well the system works if some significant fraction in its components is inactivated. Thus, we define robustness to be the average fraction of effectiveness remaining if a random 30 percent of the instances of Type D entities are removed from the system. It can be determined exhaustively by systematically removing all possible 30 percent subsets of Type D, computing and averaging the effectiveness values; or it can be determined stochastically by removing a few sample 30 percent subsets to obtain a sample set of effectiveness values.

3.2.10 Predictability

Predictability is intended to measure the degree to which the system performs as expected. As a measure, we again select one related to effectiveness. As pointed out in 3.2.1, the effectiveness is determined by finding chains of Type A entities that satisfy certain constraints. It is evident that in constructing these chains, the order in which elements are selected can affect the number that can be created before the capacity of some B-type entity or the flow rate of some C-type entity is exhausted. The effectiveness number, itself, represents the optimal order of chain construction. Other orders could produce reduced values of effectiveness. On this basis we define predictability as follows:

$$\text{Predictability} = \frac{\text{Average of effectiveness as determined by randomly selected chains}}{\text{Effectiveness as determined by an optimal chain}}$$

3.2.11 Research Performance Measures Based on System Performance

Ten quantified system PMs have been defined. We relate these to research PMs exactly as was done for development in 3.1. Let:

M be a system PM

M_c be its value for the "control" system

M_g be its value for the initial "goal" system

M_r be its value for the current revised goal system

M_n be its value for the "current state-of-the-art" system

where the items in quotes are as defined in 3.1. Then,

M_g/M_c = conceptual improvement ratio in measure M achievable through the research

M_n/M_c = improvement ratio in measure M already achieved by the research

M_r/M_n = improvement ratio in measure M still achievable through continuing the research

Thus, we have three research PMs for each system PM, or a total of 30 to add to the 12 already described in 3.1, giving a total of 42 research PMs for which values can be generated. To measure the research performance for a BMD DDP design methodology.

3.3 SCHEDULE OF EVALUATION EVENTS

The evaluation of research PMs consistent with this modeling technique would consist of the following steps:

1) Develop activity tables for two potential design methodology models:
    a) A "revised goal" model, representing the Honeywell/GRC strawman development method as envisioned at that point in time.
    b) A "current state-of-the-art" model, representing the SDS method as adjusted to incorporate features that have been validated by the research to date.

2) Compute the development times, development costs, and uncertainties in these quantities as implied by the activity tables.

3) Compute the 12 research performance measures defined in Section 3. 1.

4) Develop system model descriptions following the method of Section 3. 2 for each of the two development methodology models described above.

5) Compute the 10 system PMs for each of these system models as described in 3. 2. 1 through 3. 2. 10.

6) Compute the 30 corresponding research PMs as described in 3. 2. 11.

3. 4 ALTERNATIVE MODELING APPROACHES

A second modeling approach for performance evaluation of design methodologies is the more general approach which was proposed by Stewart (STE 76) for analysis of the effectiveness of the design process within a specific organization. This approach is based on J. Forrester's Industrial Dynamics feedback loop modeling of decisions and actions within an organization, and is concerned with achieving improvement in the efficiency and effectiveness of design decisions and actions of human designers through definition of appropriate design methodologies and automation of design techniques given a particular organizational framework and environment.

SECTION 4
LITERATURE SEARCH OF RESEARCH PERFORMANCE MEASURES

4.1 INTRODUCTION

There is a great deal of disagreement in the literature regarding the most appropriate methods for evaluating research performance. No single method is approved by any large group of experts in the field. Numerous dichotomies exist within the debate on evaluation of research. Some of the topics of disagreement are the roles of subjective versus objective evaluation techniques; quantitative versus qualitative techniques; efficiency measures versus effectiveness measures; and the appropriateness of using quantitative analysis techniques on subjective, qualitative data.

The majority of research evaluation techniques have been developed and used extensively by large business concerns in the United States. As a result, most of the pioneering work in the field is concerned with calculating the impact of in-house research on business revenue or with the development of new business opportunities. These types of analyses rely heavily on tangible, objective data, such as investment costs, time spent on research, or revenue generated by research. These quantifiable measures have been used to calculate rates of return on research, research utility, opportunity generated and cost effectiveness measures. These measures are based upon tangible data, primarily time and money.

Another set of research evaluation techniques relies upon subjective and less tangible data. These techniques include the use of expert panels and managerial opinion to create indices of performance. Although these methods are open to much criticism and outside error, they are commonly used and accepted. Many firms believe that expert panels are the only acceptable method available for evaluating nonquantifiable or intangible research outcomes. Several firms use subjective analyses together with objective analyses to evaluate the overall performance of a particular project.

The disagreement on the meaning of efficiency and effectiveness of research is highlighted by the fact that some authors treat these as two completely separate and unrelated topics while others contend that they are the same thing. The authors of the present report feel that efficiency and effectiveness are interrelated and both must be evaluated to arrive at a comprehensive evaluation of a project.

On the one hand, efficiency measures, for the most part, attempt to compare the progress of a project with the scheduled progress. It is a comparison of actual achievement with planned achievement. Several common evaluation tools are used with this type of analysis, including Management by Objectives (MBO), Program Evaluation and Review Technique (PERT), and Critical Path Management (CPM). These techniques evaluate the progress and timeliness of a research program. Timely progress, of course, is critical to a good research effort.

Effectiveness, on the other hand, attempts to relate the degree to which the research effort is meeting the goals of the research. Effectiveness measures also attempt to determine the impact of the research results on future activity.

Measuring the effectiveness of research is a highly subjective topic. No single measure of research effectiveness has been developed that can be applied uniformly to all research programs. Effectiveness measures must be tailored to fit a particular research project. The results must be interpreted with a full understanding of the evaluation procedures limitations.

As indicated in Reference 3, it is often wise to use a combination of several evaluation procedures; this permits the comparison of results and acts as a test of the evaluation outcomes.

The information that follows is a brief summary of the literature concerning the evaluation of research. Some of the more common evaluation techniques have been abstracted to provide the reader with a brief survey of the field.

The abstracts are divided into three major categories, Improvement Comparison Methods, Goals Achievement Methods, and Impact Assessment. The first category contains two techniques commonly used to compare a proposed system with one already in existence. The second category contains abstracts of three methods used to determine the extent to which research goals are being met by research being conducted. The third category abstracts a technique for identifying and assessing unintended impacts.

The categories and methods abstracted are not intended to be an exhaustive presentation of the available techniques and types of performance measures. Rather they are intended to be categories and techniques that seem appropriate for application in the Ballistic Missile Defense Advanced Technology Center (BMDATC) project.

4.2 PERFORMANCE EVALUATION MEASURES: IMPROVEMENT COMPARISON METHODS

Improvement comparison methods for evaluating research results make a comparison between the situation created by the newly developed technology and the preexisting situation, or the improvement that would have naturally occurred in the absence of the new research.

The two basic areas in which improvements are measured are the production process and the value in use. The points of tangible comparison most frequently used are:

- Cost -- to acquire or produce
- Cost -- in operation
- Time -- to produce
- Time -- savings in operation
- Size -- space savings in use
- Power -- increase in capacity

4.2.1 Tangible Comparisons

The field of engineering economics provides an evaluation of the cost of production. In the terminology of D. W. Collier and R. E. Gee (Reference 4), this is the "value-in-use" to the producer. Their explanation of the cost savings is given below:

> A value analysis starts with a detailed study of present ways of filling a function, including consideration of all cost components going into accomplishing that function. One then utilizes his new material or new idea to accomplish that same function in a different way. He must redesign, recalculate, reconsider all aspects of carrying out the function in the new way, and thoroughly cost every aspect of accomplishing the function. The costing must be over a specific time period carefully chosen to depict significant differences in product life under enduse conditions.
>
> The user's cost of carrying out his function is:

$$\text{Use-Cost} = C_m + C_a + C_s \qquad (1)$$

where

Use-Cost = Total costs, \$/time period

C_m = All direct mill costs, including materials, labor, and depreciation

C_a = All relevant allocated costs, including management expenses

C_s = All selling costs, including freight and all marketing expenses

If the use-cost associated with new technology is no greater than the use-cost associated with existing technology, a rational basis exists for the user's consideration of the new. This notion underlies the definition of Opportunity.

It should be reemphasized that value-in-use is properly used only when one is considering a function which can be defined in unambiguous terms. Value-in-use does not include consideration of any value elements not quantifiable by conventional engineering economics, such as aesthetics, are incorporated in factors influencing the conversion of Opportunity to sales (penetration). For these reasons, the applicability of value-in-use to most consumer products has not been particularly helpful; most experience has been with industrial products.

Opportunity exists when the use-cost of a new technology is less than the use-cost of its primary competition. If the primary competition is difficult to determine, the use-cost using each competitive method is calculated and the lowest use-cost method, the primary competition is that method giving the next-to-lowest use-cost. If any of the competitive methods gives a lower use-cost than that of the new method, no Opportunity is assumed in that end-use.

Value-in-use is defined as that price of the new product which makes the user's use-cost equal to his use-cost with the best alternative method. Using a prime mark to identify the user's cost elements with the new technology, the basic use-cost equation is:

$$C_m + C_a + C_s = C'_m + C'_a + C'_s \qquad (2)$$

Letting:

$$C_m = QP + C_o \qquad (3)$$

where

Q = The quantity of competitive product consumed over the specified time period

P = The unit price of the competitive material

C_o = All other direct mill costs using the competitive material

and again using prime marks to identify the new product:

$$P' = \frac{QP + (C_o' \cdot C_o') + (C_a \cdot C_a') + (C_s \cdot C_s')}{Q'} \qquad (4)$$

Potential is converted to Opportunity if the value-in-use for the new product in the specified year, equals, or exceeds the projected selling price for the stated year. For the purpose of calculating Opportunity in terms of earnings, the selling price is usually assumed to be equal to the value-in-use. While pricing at value-in-use is unrealistic in practice (the price must be set below value-in-use to provide a driving force for penetration), it is done to avoid impinging on a selling price decision, which is the province of marketing.

It should be clear that any change in methods available to the customer will lead to a change in the value-in-use of the new product. Every value-in-use figure must therefore be related to a specific period in time. Projections of value-in-use require the assessor to anticipate how the customer's technology might change and what new competitive products might be made available to him.

This method of evaluation has its primary application in industrial production. In adapting it to the present purpose of research evaluation, cost savings in production are still relevant, but can be implicitly considered in the competitive bidding process. From a military research standpoint, measurement of operating improvements may be more useful. In this area, improvements would include cost savings that are due to energy efficiency, time savings that are due to better coordination, and increased effectiveness resulting from the greater capability of the new technology. This last area is the hardest to quantify. Surely, the results are tangible, but quantifying their utility is a more subjective procedure. An attempt at dealing with this problem is described in Reference 5.

4.2.2 Intangible Index Comparisons

Research outputs that can be measured in dollars or minutes are relatively easy to assess and are considered tangibles. Many methods for evaluating tangible inputs and outputs from research have been developed and used successfully in recent years. However, the inputs and results of research are not always tangibles and therefore do not easily lend themselves to quantitative analysis. Outputs that involve comparison of characteristics or attributes such as ease of operation, robustness, maintainability, or growth capability are more abstract and require a different type of analysis.

A number of authors believe that the best way to compare intangibles in a proposed system with intangibles in an existing system is to construct an index of critical factors for each system and then compare the indices on a point-for-point basis. A number of index construction methods are commonly used by industrial firms interested in comparing one process or product with another. The most commonly used methods generally allow the comparison of an entire process with another or, at a lower level, one task with another. This permits step-by-step comparison as well as process with process comparison.

Indices are normally used with easily quantifiable data, such as cost or time. However, a number of satisfactory evaluation indices have been developed using more subjective data. The method to be discussed here uses a panel of experts and a subjective rating scale to evaluate two different processes. Each process is subdivided into a number of critical attributes. The panel of experts is then asked to rate each attribute on a scale of 1 to 4, with 4 being superior. The result, in a simple example, is an index for each attribute for each system, as shown in the matrices that follow.

Process I (Existing)

Index Matrix

Attribute	Expert:	I	II	III	IV	V	VI	Index
A		3	3	2	4	1	4	17
B		2	2	2	1	0	4	11
C		3	3	3	3	2	2	16
D		3	2	2	3	2	4	16
E		3	2	2	2	1	3	13
								73
								$\bar{X} = 14.6$

Process II (Proposed)

Index Matrix

Attribute	Expert:	I	II	III	IV	V	VI	Index
A		4	2	0	3	2	3	14
B		4	2	1	3	2	4	16
C		3	2	2	3	2	4	16
D		3	3	1	2	3	4	16
E		3	2	1	2	2	3	13
								75
								$\bar{X} = 15$

In the example cited above, the results of the analysis show that the panel of experts think that the two systems (the proposed system and the existing system) are very similar in their overall characteristics. The overall score for Process I was 73 and the score for Process II was 75. The mean scores were 14.6 and 15, respectively. A more detailed analysis shows that the experts believe that Process I is superior to Process II in Attribute A. However, Process II is superior to Process I in Attribute B. The tradeoffs between the two systems are identified through this process, allowing the existing and proposed systems can be judged accordingly.

A more complex evaluation might include subprocesses as well as critical attributes. In this case, the cumulative score of the experts is placed in the appropriate cell, as shown in the matrix that follows:

Process I

Attribute	Subprocess:	I	II	III	IV	Index
A		16	15	14	11	56
B		15	15	13	12	55
C		13	12	14	15	54
D		16	16	13	16	61
		60	58	54	54	226

$\overline{X}$ = 56.5

This more-detailed matrix would then be compared with a similar matrix for Process II. The row and column analysis would be similar to that in the first example.

This format can be expanded to three dimensions if necessary. It is also a simple matter to weight the more important subprocesses or attributes if they are particularly important. Scores may also be normalized if desired.

This type of comparison process has distinct advantages and disadvantages that must be weighed before it can be usefully applied. It is a subjective analysis and is therefore subject to error and outside influence. It also requires that a number of people be relatively familiar with the systems being compared.

An advantage to this system is that it can be quickly and easily performed using a number of experts. It permits an assessment of each subprocess and attribute as well as an overall analysis. This means that a perceived weakness in a row or column can be identified and is not buried in a massive overall evaluation. The index system also permits the researcher to weight and normalize scores where necessary, allowing an element of perspective to enter the analysis. This form of analysis can also be easily used to verify or disprove the results of another form of analysis.

The panel index comparison is an integral part of several different evaluation methods. It is frequently used as one step in a more-complex analysis; but as shown in Reference 6 it can also be used successfully independent of other processes.

4.3 PERFORMANCE EVALUATION MEASURES: GOALS ACHIEVEMENT METHODS

Improvement comparison methods are basically a comparison with history. The newly developed technology is compared with a preexisting technology on the basis of tangibles, intangibles, and anticipated outcomes.

In contrast to improvement comparison methods, goals achievement methods have a future orientation. At the outset of a research project, requirements are derived from the goals and objectives of the organization which then serve as a standard to evaluate the effectiveness of the research performed.

Goals exist on several different levels within an organization and within the larger environment of the organization. For the present purpose of research evaluation, the goals of the organization and its component parts are of prime importance. It is important that these subgoals be consistent with the larger goals of the organization. The various levels of goals within the organization and the research program can be ranked as follows:

1) Organization goals
2) Research program goals
3) Research project goals

4) Project tasks

5) Personnel goals

Beginning at the bottom of the list, the role of personnel evaluation is to assess the extent to which individuals are fulfilling organization goals. Perhaps the best known technique in this area is Management By Objectives.

4.3.1 Management By Objectives (MBO)

The basic purpose of Management By Objectives is to orient personnel toward results. In setting individual objectives, it also serves as a means of performance evaluation. In the context of research management and evaluation, the objectives on an individual level are directly related to the project, program, and organizational goals. In this sense, if the goals are internally consistent and accomplished on an individual level, then the goals of the organization should be met. Since the literature on MBO is so well known and available (References 8, 9 and 10), it is believed by the authors of this report that no further elaboration is needed.

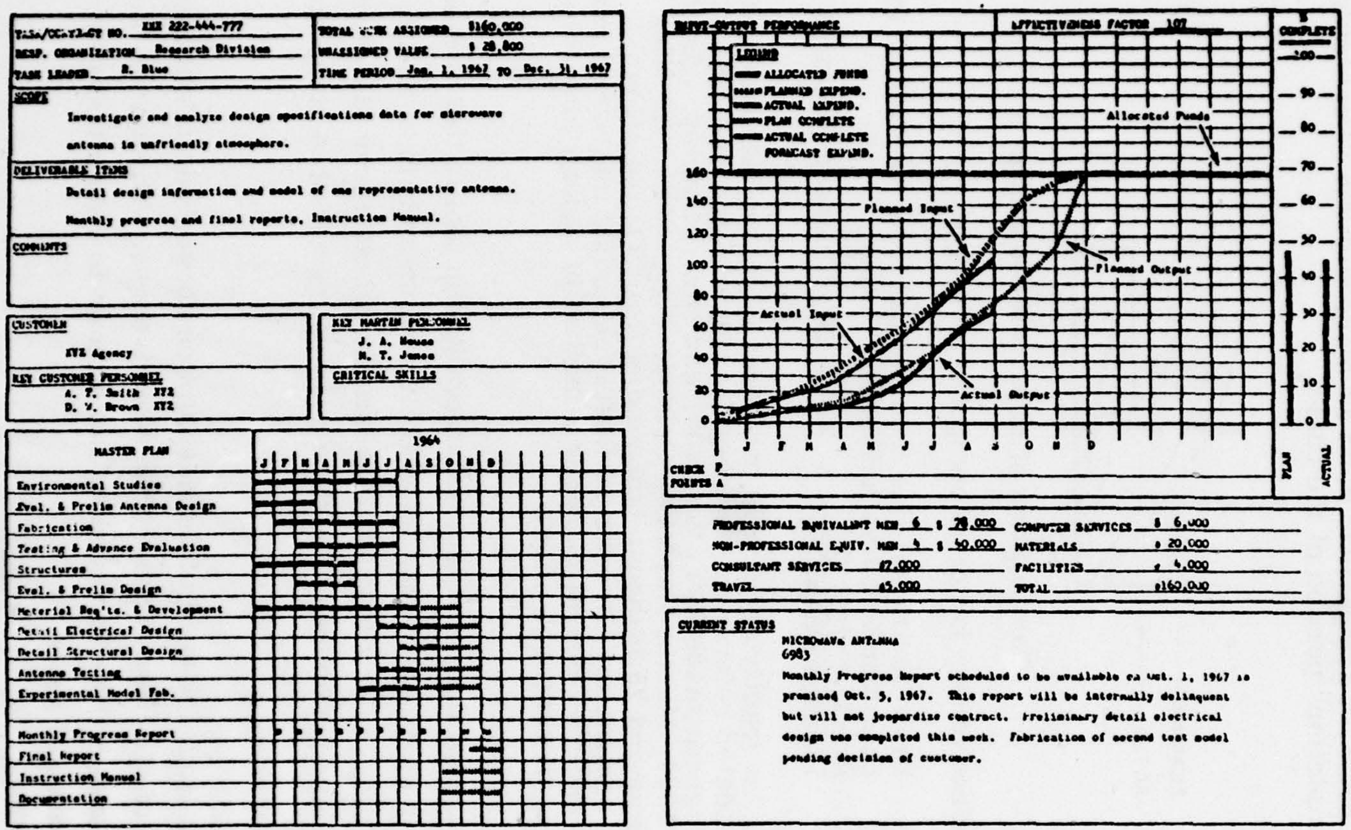
The major advantage of MBO as a research performance evaluation technique is that goals are clearly expressed at the outset of the project. In addition, because the goals are defined on an individual level, people are highly accountable for the research results.

Because MBO is basically a personnel evaluation technique and the nature of the goals established are individualized (Reference 11), the outcome of the evaluation is a judgment of individual performance. Assuming the preestablished goal was achieved, the question remains whether the solution is the best possible one or the least expensive or whether, in fact, the goal was over-achieved.

4.3.2 Input/Output Performance

Another research management technique which takes into account both cost and the effectiveness of the solution is the Input/Output Performance technique. This is basically a scheduling and review technique that provides a reading of how actual spending compares with planned spending and whether a proportional amount of the tasks have been completed. The basic form for management review is described in detail in Reference 12. It is reprinted here with an example of how the form is used:

> On the left side of the chart for management is displayed general information, such as task name, dollar value, and time parameters at the top. Immediately under this is a listing of the scope of the work to be performed, the deliverable items, and space for additional comments. Next come the customer and key customer personnel, key Martin personnel, and the critical skills required. Below this is shown the master plan by key work items plotted against time. This side of the charts tells what is to be done, by whom, for how much, where, and when.
>
> The right side of the charts shows how R&D intends to spend the resources, and how well it is following the plan. The upper dashed line is the planned dollar Input; the lower dashed line is the planned dollar Output, both plotted against time. The solid lines superimposed on the planned Input/Output lines indicate accomplishment, dollar-wise. To the right, a percent of completion versus plan is shown. Immediately below is a cost breakdown covering manpower, consulting services, travel, computer services, materials, and facilities. The last block on the chart shows the current status, which is posted weekly.
>
> To achieve a single overall measure of effectiveness, an "effectiveness factor" is computed. When the actual lines are plotted against the planned Input and Output curves, the chart may reflect many combinations of effectiveness. For example, it may show that more Input was expended than planned, and less Output was realized than planned. This certainly would indicate poor effectiveness. On the other hand, it is quite likely that for less Input than planned, the right amount of Output was accomplished. This would indicate good performance. The effectiveness factor reflects actual work accomplished for money spent and does not reflect schedule status. To establish the relationship



TASK/CONTRACT NO. XXX 222-444-777	TOTAL WORK ASSIGNED $160,000
RESP. ORGANIZATION Research Division	UNASSIGNED VALUE $ 28,800
TASK LEADER R. Blue	TIME PERIOD Jan. 1, 1967 TO Dec. 31, 1967

SCOPE

Investigate and analyze design specifications data for microwave antenna in unfriendly atmosphere.

DELIVERABLE ITEMS

Detail design information and model of one representative antenna.

Monthly progress and final reports, Instruction Manual.

COMMENTS

CUSTOMER

XYZ Agency

KEY CUSTOMER PERSONNEL

A. T. Smith XYZ
D. W. Brown XYZ

KEY MARTIN PERSONNEL

J. A. House
H. T. Jones

CRITICAL SKILLS

INPUT-OUTPUT PERFORMANCE

EFFECTIVENESS FACTOR 107

PROFESSIONAL EQUIVALENT MEN 6 $ 78,000	COMPUTER SERVICES $ 6,000
NON-PROFESSIONAL EQUIV. MEN 4 $ 40,000	MATERIALS $ 20,000
CONSULTANT SERVICES $7,000	FACILITIES $ 4,000
TRAVEL $5,000	TOTAL $160,000

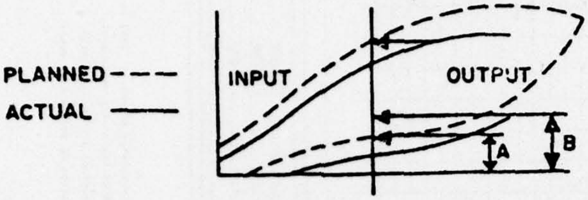
CURRENT STATUS

MICROWAVE ANTENNA
6983

Monthly Progress Report scheduled to be available on Oct. 1, 1967 is promised Oct. 5, 1967. This report will be internally delinquent but will not jeopardize contract. Preliminary detail electrical design was completed this week. Fabrication of second test model pending decision of customer.

Input/Output Performance Chart

to schedule, a comparison must be made with the "Percentage of Completion" portion of the chart.



Computation of effectiveness factor

The effectiveness factor is computed as follows:

1) Draw horizontal line from Actual Input to Planned Input.
2) Drop vertical line to Planned Output.
3) $\frac{\text{Actual Output (B)}}{\text{Planned Output (A)}} \times 100 = \text{Effectiveness Factor \%}$

As can be readily seen, this one chart gives top management a consolidated reference, showing the requirements of the task, the progress made from the start, and current status. As of early 1968, Martin Marietta, Orlando Div., was controlling 30 contracts and 75 in-house tasks using Input/Output charts. These provide the Director of Research, other interested executives, and the General Manager a method of rapidly assessing the overall research effort.

The advantage of this technique as an evaluation tool is that it measures the output in relation to input as well as measuring both in relation to the time schedule. This allows an assessment of both above standard and below standard performance in time and cost required to complete the project.

Despite the language used in the description above, this method provides a measure of efficiency rather than effectiveness. Though the method measures task completion, it does not attempt to relate tasks to research requirements and goals.

4.3.3 Project Goals Achievement

One method for determining the degree to which research goals are being met is the Project Goals Achievement method developed at Borg-Warner (Reference 6). This evaluation method and variations of it have gained considerable popularity in recent years. It is normally used to evaluate a number of projects over time, but can easily be adapted to the evaluation of a single large project.

The system requires that the client and contractor (together) evaluate the project in question at specified time intervals. The client and contractor form a panel that rates each task or subtask of the project with regard to the degree to which the objective of the task or subtask has been met, given the dollar and time expenditure. Each task or subtask is ranked on a scale of 0 to 4. The resulting score is then multiplied by the money spent on the task to date, and then divided by the total expense of the project. The result is a weighted index of the performance on a task versus its objectives.

This process is useful in evaluating large projects that run for a period of a year or more. Periodic evaluations provide an overall picture of how well the research is meeting its objectives and whether or not there has been an improvement or decline in performance. An added benefit of this process is that it forces the client and contractor to meet and agree on objectives and progress on a regular basis. Timely mid-course corrections may result from these meetings.

If a project is evaluated in this fashion over a period of time, an evaluation matrix (see example below) is developed. The matrix allows easy analysis of the progress being made in each area of research and permits upper level management to keep close tabs on research without great time expenditure.

Sample Project Goals Achievement Matrix

Task 1

	1/1/75	1/6/75	1/1/76	1/6/76
Subtask A	0.4	0.9	0.9	1.6
B	0.3	0.6	0.9	1.4
C	0.0	0.5	1.1	1.9
D	0.7	0.9	1.6	2.1
E	0.6	1.1	1.7	2.3
	0.35	0.80	1.25	1.86

In the preceding illustration, an implicit evaluation of goals achievement was made in the rating of tasks and subtasks; the illustration of this process can be clarified by rating tasks with respect to the accomplishment of specific goals. This is a useful method of evaluating the effectiveness of the research after the work has been done as well as a means of ensuring that the tasks undertaken actually fulfill the goals and requirements of the research. In the example given below, each task is explicitly evaluated and rated for its contribution to a list of research objectives or goals:

	Task 1	Task 2	Task 3	Task 4	Task 5	Total
Goal A	0.25	0.05	0	0.55	0.15	1.00
B	0	0.10	0.80	0	0.10	1.00
C	0.10	0	0	0.50	0.30	0.90
D	0.35	0.40	0.10	0	0.10	0.95

The advantage of this method is that it not only assesses the extent of achievement of each goal, but also shows how much each task contributed to that achievement. This allows a weighting of the importance of the tasks, which is useful in setting priorities. The disadvantage of this method is that it does not differentiate levels of value for the different goals, but assumes that all are of equal value. This is certainly not the case. An attempt at quantifying the value of research objectives for military applications can be made using indifference analysis. An example of this technique is provided in Reference 5, on quantifying military utility.

4.4 PERFORMANCE EVALUATION MEASURES: IMPACT ASSESSMENT

The Goals Achievement Methods outlined in Section 4.3 are intended to measure the extent to which established goals are achieved. Equally important, and often harder to accomplish, is the identification and assessment of unintended affects. Impact Assessment attempts to meet this challenge in the broadest possible way.

4.4.1 Technology Assessment

The definition of Technology Assessment (TA) in current use was articulated (Reference 13) by J. Coates, former manager of the TA program at the National Science Foundation:

> Technology Assessment is a class of policy studies which systematically examines the effects on society that may occur when a technology is introduced, extended, or modified with special emphasis on those consequences that are unintended, indirect, or delayed .... Comprehensive impact or assessment studies are a class of holistic studies which attempt in some sense to embrace everything that is important with regard to a technology .... One characteristic of holistic thinking is that we do not know how to do it routinely; secondly, it almost certainly cannot be done routinely; and, thirdly, it is not a scientific or an engineering or a disciplinary enterprise. It is essentially an art form.

The analytical parameters of TA are very broad. To paraphrase S. R. Arnstein (Reference 4), comprehensive technology assessment provides an evaluation of technical efficacy, economic feasibility, safety risks, and public policy options, including second- and higher-order effects on all the relevant impact domains, interested parties, and social systems.

Technology Assessment is not an identifiable technique of evaluation; rather it is an approach to evaluation which uses several methods of evaluation depending on the type of analysis required. For example, analytical tools might include cross impact analysis, dynamic modeling, factor analysis, opinion surveys, relevance trees, futures scenarios, sensitivity analysis, and input/output analysis.

The evaluation is generally made by a multidisciplinary team of experts drawn from fields relevant to the technology. The things which distinguish TA are the holistic approach to problems, and a focus on identifying unintended impacts of the proposed technology.

One limitation of TA is that the areas of social impact in which it deals are often difficult to quantify. In addition, many studies done in the past have suffered from unclear definition of the area of assessment.

The main advantage of using a comprehensive TA approach is the potential for identifying unintended adverse impacts, and suggesting alternative paths of development.

4.5 SUMMARY AND CONCLUSIONS

Several research evaluation techniques are currently being used successfully. The majority have been developed by major business concerns as tools for evaluating their own in-house research efforts. Nearly every published technique is designed to evaluate a specific type of research in a specific form. As a result, the published techniques are applicable only under very specific circumstances and do not lend themselves to use elsewhere.

Although the specific evaluation tool cannot be lifted from its context and applied elsewhere, the basic underlying theories of the published techniques can be used to design appropriate techniques for almost any application. To evaluate a specific research task, it is necessary to carefully tailor an appropriate method. We believe it is advisable to use several techniques in consort to provide verification of results and ensure a broad evaluation.

After an evaluation measure is carefully developed and applied, it is important that the results be interpreted with a full knowledge of the evaluation method's limitations. Many evaluators tend to overstate the implications of their evaluation methods. This problem can be partially overcome through the use of multiple evaluation methods.

We have presented a variety of techniques, some of which are applicable to various areas of distributed data processing research. An attempt has been made to identify the advantages and disadvantages of the various methods.

SECTION 5
PERFORMANCE MEASURES RECOMMENDATIONS

5.1 APPLICABILITY OF EVALUATION MEASURES (general)

The Improvement Comparison Methods of evaluating research are most appropriately applied to research whose output is a system or process that can be directly compared with an existing system or process. The proposed system or process, such as DDP or a specific computer architecture, is compared with the existing system on a number of characteristics, both tangible and intangible. This type of measurement is not a valid method for evaluating tasks whose results are not an alternative process or system. For instance, it is possible to compare the efficiency of one data-processing system with another; but it is not possible to quantitatively compare investigative research, such as research on design methodology or issues associated with software development.

To evaluate research which supplies such information rather than developing a process or system, a Goals Achievement Method must be employed. A Goals Achievement Method is capable of determining whether a research effort has successfully met its intended objectives. It is possible to apply both a Goals Achievement Method and an Improvement Comparison Method to a research project whose output is a system or process. This permits the evaluator to determine whether or not a proposed system is better than an existing system and also the degree to which the research has attained its intended goals. It is not always necessary or desirable to use both methods; however, in some cases it may be appropriate.

The function of Impact Assessment is to identify unintended effects and suggest alternative paths of development. Wherever it is feasible, this type of analysis is of value because it in some measure ensures that our solutions do not generate a new set of problems that are greater than those we solve.

In selecting research evaluation methods for DDP research, it is necessary to divide the program into its various tasks and subtasks. If the intended result of a task or subtask is purely to develop methodology, a Goals Achievement Method should be applied, as discussed in detail in Section 3. Generally, if the intended output is a system or process, then an Improvement Comparison Method is applied. In the case of DDP research, a combination of methods is indicated.

Table 4 indicates the type or types of evaluation methods that might be applied to DDP research. Applicability of evaluation measures is indicated for each task according to its general nature. The subtasks within a task involved may often require the use of differing techniques.

TABLE 4. APPLICABILITY OF MEASURES TO THE MAJOR RESEARCH ELEMENTS OF THE BMDATC PROJECT

Major Research Element	Improvement Comparison Methods	Goals Achievement Methods
Technology Development :		
• Distributed computer architecture	X	X
• System algorithms	X	X
• Computer hardware	X	X
Technology Application:		
• Develop methods and procedures		X
• Applying DDP to BMD		X
• Feasibility analysis		X
Quantifying Payoff:		
• Of DDP over centralized processing	X	
Technology Evaluation:		
• Of DDP concepts	X	X
• For BMD applications	X	X

5.2 RECOMMENDATIONS CONCERNING APPLICABILITY OF EVALUATION MEASURES TO DDP PHASE I

Because this research was primarily oriented toward concept definition and exploration, quantitative measures (except for simple brute-force measures like number of architectures evaluated, number of payoffs quantified, number of architectures defined, number of experiments defined, number of different methodologies developed, etc.) cannot be effectively applied at this stage of the research.

Pursuit of the experiment and simulation techniques described in Volume VII will provide quantitative measurements of research effectiveness during phase 2 of the program, and, as the research recommended in Volumes II through VII and IX is pursued, the evaluation measures described in Section 4 of this volume can be selectively applied. For evaluation of performance of this concept definition phase, however, we recommend that qualitative criteria be applied and the index of critical factors be applied to evaluate the overall program performance.

Suggested evaluation factors are:

1) Have requirements of the statement of work been met?
2) Have new disciplines been applied to DDP research?
3) Has the definition of DDP been clarified?
4) Have the critical issues associated with the application of DDP to BMD been identified?
5) Has an approach and plan for resolution of these issues been defined?
6) Has the research added to the state of the art on DDP?
7) Have the payoffs of DDP versus CDP been demonstrated and quantified?

8) Have new design methodologies been defined?

9) Has a systems approach to DDP design for BMD been considered, as well as a technology approach?

10) Have the requirements for the output of ARE been identified?

11) Has a feasible approach for demonstration and experimentation of DDP for BMD been defined?

These are the major criteria which could be used to apply the index of critical factors performance measurement approach.

REFERENCES

1. West, J. D., and Levy, F. K., A Management Guide to PERT/CPM, Prentice-Hall, Englewood Cliffs, NJ, 1969.
2. Data Processing Subsystem Engineering Benchmark Experiment, General Research Corporation Report CR-3-685, March 1976.
3. Lipetz, Ben-Ami, The Measurement of Efficiency of Scientific Research, Intermedia Inc., Carlisle, Massachusetts, 1965.
4. Collier, Dr. Donald W., Borg-Warner Corporation, and Gee, Robert E., duPont de Nemours and Company, Inc. Article reprinted in Research Management, Industrial Research Institute, New York, May 1973.
5. Cetron, Marvin J., "Quantifying Military Utility," Technical Resource Management, Cambridge, Massachusetts, 1969. See Appendix: "A Survey of Quantitative Methods."
6. Collier, Donald W., "Measuring the Performance of R&D Departments," Research Management, Industrial Research Institute, New York, March 1977.
7. Tauss, Kurt H., "A Pragmatic Approach to Evaluating R&D Programs," Research Management, Industrial Research Institute, September 1975.
8. Odiorne, G., Management By Objectives--A System of Managerial Leadership, Pitman, 1965.
9. Hughes, C. L., Goal Setting: Key to Individual and Organizational Effectiveness, American Management Association, 1965.
10. Drucker, Peter R., "Controls, Control and Management," Management Controls: New Directions in Basic Research, McGraw-Hill, 1964.
11. Levinson, 1970.
12. Soistman, E. C. (Director of program planning Orlando Division Martin Marietta Corporation, Orlando, Florida.) Article reprinted in the Handbook of Industrial Research Management, Reinhold Book Corporation: New York, 1968.
13. Coates, J. Articles reprinted in Institute of Electrical and Electronics Engineers (IEEE) Transactions on Systems, Man and Cybernetics, Vol. 7, No. 8, August 1977.
14. Arnstein, S. R., "Technology Assessment: Opportunities and Obstacles," Institute of Electrical and Electronics Engineers (IEEE) Transactions, August 1977.

BIBLIOGRAPHY

Collier, Donald W., and Gee, Robert E., "A Simple Approach to Post-Evaluation of Research," Research Management, May 1973.

Gee, Robert E., "The Opportunity Criterion--A New Approach," Research Management, May 1972.

Heyel, Carl, Handbook of Industrial Research Management, Reinhold Book Corporation, 1968.

Quinn, James B., Yardsticks for Industrial Research, Ronald Press, 1959.

Roman, Daniel D., Research and Development Management, Meredith Corporation, New York, 1968.

Stewart, Barbara C., "Design Analysis: An Approach for Improving the System Design Process," UCLA Computer Science Department, Los Angeles, December, 1976.