LEVEL II

DEPARTMENT OF COMMERCE · NATIONAL BUREAU OF STANDARDS

NBS TECHNICAL NOTE 890

U.S. DEPARTMENT OF COMMERCE / National Bureau of Standards

AD A106487

# PRODUCTIVITY MEASUREMENT IN R&D:

## Productivity Measurement Experiment (PROMEX) in Selected Research and Development Programs at the National Bureau of Standards

# NATIONAL BUREAU OF STANDARDS

The National Bureau of Standards[1] was established by an act of Congress March 3, 1901. The Bureau's overall goal is to strengthen and advance the Nation's science and technology and facilitate their effective application for public benefit. To this end, the Bureau conducts research and provides: (1) a basis for the Nation's physical measurement system, (2) scientific and technological services for industry and government, (3) a technical basis for equity in trade, and (4) technical services to promote public safety. The Bureau consists of the Institute for Basic Standards, the Institute for Materials Research, the Institute for Applied Technology, the Institute for Computer Sciences and Technology, and the Office for Information Programs.

**THE INSTITUTE FOR BASIC STANDARDS** provides the central basis within the United States of a complete and consistent system of physical measurement; coordinates that system with measurement systems of other nations; and furnishes essential services leading to accurate and uniform physical measurements throughout the Nation's scientific community, industry, and commerce. The Institute consists of the Office of Measurement Services, the Office of Radiation Measurement and the following Center and divisions:

Applied Mathematics — Electricity — Mechanics — Heat — Optical Physics — Center for Radiation Research: Nuclear Sciences; Applied Radiation — Laboratory Astrophysics[2] — Cryogenics[2] — Electromagnetics[2] — Time and Frequency[2]

**THE INSTITUTE FOR MATERIALS RESEARCH** conducts materials research leading to improved methods of measurement, standards, and data on the properties of well-characterized materials needed by industry, commerce, educational institutions, and Government; provides advisory and research services to other Government agencies; and develops, produces, and distributes standard reference materials. The Institute consists of the Office of Standard Reference Materials, the Office of Air and Water Measurement, and the following divisions:

Analytical Chemistry — Polymers — Metallurgy — Inorganic Materials — Reactor Radiation — Physical Chemistry.

**THE INSTITUTE FOR APPLIED TECHNOLOGY** provides technical services to promote the use of available technology and to facilitate technological innovation in industry and Government; cooperates with public and private organizations leading to the development of technological standards (including mandatory safety standards), codes and methods of test; and provides technical advice and services to Government agencies upon request. The Institute consists of the following divisions and Centers:

Standards Application and Analysis — Electronic Technology — Center for Consumer Product Technology: Product Systems Analysis; Product Engineering — Center for Building Technology: Structures, Materials, and Life Safety; Building Environment; Technical Evaluation and Application — Center for Fire Research: Fire Science; Fire Safety Engineering.

**THE INSTITUTE FOR COMPUTER SCIENCES AND TECHNOLOGY** conducts research and provides technical services designed to aid Government agencies in improving cost effectiveness in the conduct of their programs through the selection, acquisition, and effective utilization of automatic data processing equipment; and serves as the principal focus within the executive branch for the development of Federal standards for automatic data processing equipment, techniques, and computer languages. The Institute consists of the following divisions:

Computer Services — Systems and Software — Computer Systems Engineering — Information Technology.

**THE OFFICE FOR INFORMATION PROGRAMS** promotes optimum dissemination and accessibility of scientific information generated within NBS and other agencies of the Federal Government; promotes the development of the National Standard Reference Data System and a system of information analysis centers dealing with the broader aspects of the National Measurement System; provides appropriate services to ensure that the NBS staff has optimum accessibility to the scientific information of the world. The Office consists of the following organizational units:

Office of Standard Reference Data — Office of Information Activities — Office of Technical Publications — Library — Office of International Relations — Office of International Standards.

---

[1] Headquarters and Laboratories at Gaithersburg, Maryland, unless otherwise noted; mailing address Washington, D.C. 20234.

[2] Located at Boulder, Colorado 80302.

# PRODUCTIVITY MEASUREMENT IN R&D:

## Productivity Measurement Experiment (PROMEX) in Selected Research and Development Programs at the National Bureau of Standards

John T. Hall

Office of the Associate Director for Administration
National Bureau of Standards
Washington, D.C. 20234

and

Roger A. Dixon

Institute for Basic Standards
National Bureau of Standards
Boulder, Colorado 80302





U.S. DEPARTMENT OF COMMERCE, Rogers C. B. Mo .on, *Secretary*
James A. Baker, III, *Under Secretary*
Dr. Betsy Ancker-Johnson, *Assistant Secretary for Science and Technology*
NATIONAL BUREAU OF STANDARDS, Ernest Ambler, *Acting Director*

Issued December 1975

National Bureau of Standards Technical Note 890

Nat. Bur. Stand. (U.S.), Tech. Note 890, 52 pages (Dec. 1975)

CODEN: NBTNAE

U.S. GOVERNMENT PRINTING OFFICE
WASHINGTON: 1975

For sale by the Superintendent of Documents, U.S. Government Printing Office, Washington, D.C. 20402
(Order by SD Catalog No. C13.46:890). Price $1.10 (Add 25 percent additional for other than U.S. mailing).

Contents


Abstract . . . . . . . . . . . . . . . . . . . . . . . . 1

Statement of Purpose . . . . . . . . . . . . . . . . . . 2

Summary . . . . . . . . . . . . . . . . . . . . . . . . 3

I. Overview of Experiment . . . . . . . . . . . . . . . 5
- 1. Organization of Report . . . . . . . . . . . . . 5
- 2. Background . . . . . . . . . . . . . . . . . . . 5
- 3. Concepts and Rationale . . . . . . . . . . . . . 6

II. Conduct of Experiment . . . . . . . . . . . . . . . 8
- 1. Interviews . . . . . . . . . . . . . . . . . . . 8
- 2. The Process . . . . . . . . . . . . . . . . . . 8
- 3. Inputs to the Process . . . . . . . . . . . . . 13
- 4. Value Analysis . . . . . . . . . . . . . . . . . 17
- 5. Outputs of the Process . . . . . . . . . . . . . 20
- 6. Impact of the Process. . . . . . . . . . . . . . 22
- 7. Communications Study . . . . . . . . . . . . . . 22

III. Results . . . . . . . . . . . . . . . . . . . . . 27
- 1. Examination of Findings . . . . . . . . . . . . 27
- 2. Conclusions . . . . . . . . . . . . . . . . . . 30
- 3. Recommendations . . . . . . . . . . . . . . . . 32

Bibliography . . . . . . . . . . . . . . . . . . . . . 34

Figures
- 1. Dimensions of Productivity . . . . . . . . . . . 7
- 2. Promex Continuous Observations . . . . . . . . . 12
- 3. Work Sampling Summary. . . . . . . . . . . . . . 14
- 4. Promex Process Outline Chart . . . . . . . . . . 15
- 5. Estimated Manyears & Costs by Program . . . . . 16
- 6. Actual Costs for Labor, Overhead, and Other Objects for FY '73 by Program . . . . . . . 16
- 7. Criteria for Evaluation: Value Analysis . . . 19

Appendices (Available on request from Management and Organization Division, NBS)
- A. Prospectus
- B. Blueprint
- C. Work sampling study
- D. Communications study
- E. Statistical analysis of rating results
- F. Output analysis: publications
- G. Report of Interviews
- H. Value Analysis

Productivity Measurement in R&D

A Report on a Productivity Measurement Experiment (PROMEX) in Selected Research and Development Programs at the National Bureau of Standards

John T. Hall
National Bureau of Standards
Washington, DC

Roger A. Dixon
National Bureau of Standards
Boulder, Colorado

*This report describes an experiment in productivity measurement conducted at the National Bureau of Standards. The experiment concludes that no matter how sophisticated the analysis and synthesis processes become, statistical counts of output media (e.g., publications, citations, invited talks) will not serve as reliable measures of R&D productivity.*

*The conduct of the experiment included a work sampling study, a communications study, an output analysis, a value analysis approach to developing criteria for selection and evaluation of programs, construct of a rating system for evaluation of programs, and construction of a model of the R&D process.*

*Key words: Impact; objectives; output; performance measurement; production indices; production measurement.*

## Statement of Purpose

The purpose of this experiment was to develop a statement on the usefulness and limitations of present and new productivity measurement techniques as applied to R&D activities of the type performed by NBS.

The experiment was conducted by staff members of the Office of the Deputy Director, IBS/Boulder, and members of the Management and Organization Division in Gaithersburg, Maryland. Participating directly in the experiment were programs from the Quantum Electronics, Time and Frequency, and Cryogenics Divisions in Boulder, Colorado. The study was coordinated by the Management and Organization Division.

### Principal PROMEX Contributors

Study Team:
- J. T. Hall, Coordinator
- J. C. Aubele
- S. Damours
- W. B. Davis
- R. A. Dixon
- I. M. Lloyd
- R. Oman

Value Analysis Group:
- B. W. Birmingham, Chairman
- J. A. Barnes
- H. S. Boyne
- R. D. Harrington
- A. R. Hauler
- R. H. Kropschot
- R. C. Sangster
- S. J. Smith

Program Managers:
- E. D. West, Laser Parameter Measurements
- H. Hellwig, Quantum Electronics Frequency Standards
- R. A. Kamper, Cryoelectronics
- D. B. Mann, Cryogenic Metrology

Other Contributors:
- T. Dillon
- S. Dunaway
- J. Ellermeier
- K. T. Higgins
- R. N. McKee
- R. L. Powell
- T. Russell
- S. L. Stevenson

Summary

This report describes an experiment in productivity measurement (PROMEX) conducted in four program areas in the Institute for Basic Standards of the National Bureau of Standards. The four program areas, located in the Bureau's laboratories in Boulder, Colorado, were:

| Program Area | Division |
|---|---|
| Laser parameter measurements | Quantum Electronics |
| Quantum electronics frequency standards | Time and Frequency |
| Cryoelectronics | Cryogenics |
| Cryogenic metrology | Cryogenics |

The conduct of the experiment included a work sampling study, a communications study, an output analysis, and value analysis approach to developing criteria for selection and evaluation of programs, construct of a rating system for evaluation of programs, and construction of a model of the R&D process.

Findings from the value analysis approach to developing and using criteria for selection and evaluation of programs included:

- value analysis can be used in the selection and evaluation of programs;
- one set of criteria can be used for both selection and evaluation;
- weighting the criteria which were developed did not in practice significantly affect the ranking of programs.

The study found that no matter how sophisticated the analysis and synthesis processes become, statistical counts of output media (e.g., publications, citations, invited talks) will not serve as reliable measures of productivity.

Recommendations include:

1. that NBS adopt a formal, open system for rating and ranking program initiatives and on-going programs;
2. that further investigation be conducted at NBS on the use of two techniques which may provide useful productivity indicators:
   - value analysis based on criteria for selection of programs and evaluation of results
   - a survey technique which focuses on specific uses of outputs;

3. that a study be conducted at NBS of certain behavioral and managerial factors involved in the R&D process to assess the impact of these factors on productivity;

4. that further work process studies be conducted in other selected areas of NBS.

# I. OVERVIEW OF EXPERIMENT

## 1. Organization of Report

This report is organized in three parts: Overview of Experiment, Conduct of Experiment, and Results. Part I, Overview of Experiment, describes the historical and environmental milieu in which the study developed and was conducted, as well as the basic concepts and rationale. Each of the sections in Part II, Conduct of Experiment, discusses a particular aspect of the study. Part III, Results, covers an examination of findings, conclusions, and recommendations.

The body of the report, consisting of Parts II and III, is organized around the PROMEX Process Outline Chart (Model) which is described in Part II, Section 2.c., Representing the Process. The appendices contain documents, studies, and reports which were produced during the course of the experiment.

## 2. Background

On several occasions over the past ten years NBS has received requests for quantifiable measures of output. Some of these requests have been for specific measures, such as units produced or unit costs. In general, the Bureau has responded to these requests with statements to the effect that it was unable to supply quantifiable measures.

In the early 1960s, a very strong effort was made by the Department of Commerce to institute work measurement for all programs in all Bureaus. NBS technical and administrative divisions were pressured to identify countable outputs. In the technical divisions, very few could even be identified. The DOC work measurement expert was unable to identify countable outputs in the technical programs or to suggest feasible ways to collect data on those tentatively identified. After about a year, the DOC work measurement effort was abandoned.

Two exploratory efforts dealing, respectively, with the output and value of NBS programs were "Some Aggregate Measures of Productivity and Determinants of Funding at NBS" (May, 1970) by Rosalie T. Ruegg and "Exploratory NBS Studies of Benefit-Cost Measurement in Research and Development" (1968) by John T. Yates, Jr., Howard E. Morgan, and Robert D. Huntoon. These studies focus on two significant concepts which are treated in the PROMEX study; output and impact.

A September 1970, letter from Senator William Proxmire, Chairman of the Joint Economic Committee, to Elmer Staats, Comptroller General, regarding the need for productivity measures in the Federal government served as the impetus for a joint Civil Service Commission, General Accounting Office, and the Office of Management and Budget productivity measurement project.

The joint CSC/GAO/OMB continuing project is the most recent attempt by the Federal government to measure productivity. The motivation for the NBS' productivity measurement experiment was the recognition of the fact that the Bureau could not provide meaningful productivity measures of R & D program activity for the Joint Committee and that R & D productivity measurement was a problem confronting many Federal agencies.

Once the decision to conduct a productivity measurement experiment at NBS was made, a "Prospectus," (April, 1972- Appendix A) providing a general overview, and a "Blueprint," (July, 1972- Appendix B) suggesting some of the details of the proposed project, were developed. Discussions of PROMEX centering around these reference documents were held with staff members from the Management and Organization Division in Gaithersburg, the Cryogenics, Quantum Electronics, and Time and Frequency Divisions in Boulder, and with representatives from the Office of Telecommunications and the National Oceanic and Atmospheric Administration in Boulder. During the early stages of PROMEX, OT and NOAA decided not to participate further. Four programs from the three Boulder Divisions were selected to participate and have done so for the duration of the PROMEX study.

## 3. Concepts and Rationale

The attempts of other Federal Departments to develop work measurement reporting systems have relied quite heavily on the more traditional measures and concepts of performance measurement. There has been little experimentation with the concept of productivity measures including the application and limitations of the concept to research programs. Most of the past efforts in this field, particularly those originating at organizational levels above NBS, did not lend themselves to experimentation and were concerned with only one or another of the many factors of productivity measurement.

Past efforts have been for the most part concerned with conveniently available output measures without regard for their quality as productivity indicators. Such efforts have for a variety of reasons relied almost completely on techniques used for planning and measuring the efficiency of manual work. The application of these techniques by organizations that rely heavily on knowledge work for their outputs has been largely one of conformance rather than real use.

It seemed a proper role for NBS to conduct a comprehensive experiment on the question of productivity measures in an R & D organization. A plan for such an experiment was developed around a definition of productivity measurement that included the classification of work, methods of measuring performance as well as effectiveness, criteria for selecting work, plan vs. accomplishment, units produced, and impact of work. Figure 1, Dimensions of Productivity, represents this definition of productivity.

The plan was to experiment with the application of these concepts to research programs and by so doing add to the basic knowledge concerning the application of the concepts of productivity measures in research organizations.

| PHASE | | MEASUREMENT FACTOR |
|---|---|---|
| *Inputs* | → | *Value of Input* |
| *Process* | → | *Efficiency in generation of output* |
| *Output* | → | *Value of output* { —*Volume of output* —*Quality of output* —*Utility of output (impact)* |

*Premise: To measure productivity in R&D operations it is necessary to take into account the measurement factors shown above.*

Figure 1, Dimensions of Productivity

The primary purpose of the proposed experiment was to enable the Bureau to speak knowledgeably on the usefulness and limitations of measurement techniques as applied to the R & D activities which we perform.

Also, it was anticipated that the experiment would help to:

- optimize resource allocation;
- sharpen criteria for technical project selection and evaluation;
- better understand the NBS decision-making processes;
- categorize the kinds of projects which should not compete with each other (e.g., pure research projects vs. research projects which have a highly visible impact on national needs).

## II. CONDUCT OF EXPERIMENT

### 1. Interviews

The first phase of the productivity measurement experiment included interviews with members of the technical staff to gather preliminary data. The interviews, conducted in September through October 1972, consisted of a comprehensive interview with each program manager and a more limited interview with professional and technical staff members working on programs included in the experiment. Those interviewed were asked for their opinions about aspects of productivity as applied to their programs, including planning, work process, effectiveness, efficiency, outputs, value analysis, impacts, and communications. A total of 37 people, including five program managers, were interviewed.

The interview results included a large amount of information on plans and work processes and relatively little information on value analysis and communications. Analysis of the interview results provided information of interest in itself, such as: a list of impediments to efficiency, a preliminary list of outputs from each of the programs, and a preliminary list of impacts for each program.

The interviews also provided the basis for further structuring of data collection for the various segments of the experiment, such as: an expressed interest in how time was actually used in the work process, and several suggested criteria for program selection and evaluation which were used as input to the value analysis segment of the study. More generally, the interviews provided the PROMEX team with the familiarity with each of the programs necessary to proceed with the whole effort, and served to introduce the team to the members of the technical staff who would be involved in the study. For further detail on the interviews, see Appendix G, "Report of Interviews."

### 2. The Process

A. Why the Process was Studied.
B. How the Process was Studied.
C. Representing the Process.

A. <u>Why the Process Was Studied.</u>

In conventional productivity measurement, where the product is reasonably straightforward and easily identified, productivity can be measured by counting units of input to the system, i.e., costs, and units of output from it, i.e., products, and comparing them by means of a ratio. The production process between inputs and outputs can be treated as a "black box" whose inner workings need only be known to the manager and staff. Productivity per se can be monitored by a person who lacks understanding of what happens inside the "black box," simply by observing the trend of the ratio from a base period of time.

In the case of measuring research productivity, the problem is much more difficult. Whereas the products of a manufacturing plant are characterized by their similarity to one another in nature and use and are therefore meaningfully counted, the obviously countable "outputs" of research, such as publications and prototype devices, are so different from one another in their content and probable impact as to render mere counting meaningless. Furthermore, tentative problem solutions and new insights into problems may be communicated to potential users and applied on a pilot basis before the formal research is completed and a publication or "output" is produced.

In a manufacturing process the causal linkages between input and output (e.g., the details of the factory production line) are consciously designed before the process is put into operation. Although unexpected problems do occur and creative solutions are worked out, the general picture of the causal chain from input to output is reasonably well understood. It is therefore feasible to trace variations in productivity to their causes in a defined process. It is also possible to predict how a change in the level of input will affect the level of output.

In research, by contrast, the process is far less predictable and controllable. Conscious overall design is less prominent, though present and necessary, and the functions of creative problem solving and innovative decision making during the process are far more prominent. As a result of this exploratory quality of the process, the set of causal linkages between inputs and outputs or results is far less obvious and predictable than in a repetitive process. Without some understanding of these causal linkages one cannot say with confidence that increasing or decreasing the resource inputs by amount "x" will produce result "y". For instance, will doubling the input dollars result in twice the quantity and/or value of the output? If this kind of statement cannot be made, can some other useful statements be made about the relationship between inputs and outputs?

Clearly the process of research is carried out by scientists and technicians in laboratories and offices doing some kind of work. The work is carried out over a period of time, and the people doing it are paid for the time they spend. How do they spend their time? What observable and measurable things happen in the process? It was to answer such questions as these that a study of the research process was included in the Productivity Measurement Experiment.

B. <u>How the Process Was Studied</u>.

In order to understand the work process and the causal linkages between inputs and results, the PROMEX team used two approaches: interviews and direct observations. The interview data were collected as a part of the comprehensive interview of all study participants conducted at the beginning of the experiment. These interviews covered other areas of the study as well as work process. Direct observation of the process was carried out using two techniques: continuous observation and work sampling.

The data collected by these two direct observation techniques are summarized in several forms including tables and a process flow diagram (see figures 2, 3, and 4).

(1) Interview data on work process

The interview questions bearing on the work process were closely related to the questions on the plans for carrying out the program. Questions on these subjects elicited responses on the content of the research and on who was assigned responsibility for what part of the future work. When supervisors were asked to describe "major phases" of the research and subdivide major phases into "activities" and "tasks", they usually spoke of each individual as being assigned a major phase of the work running concurrently with other phases carried out by other individuals. They did not conceive a phase as one step in a sequential series of steps. They did, however, subdivide each of the major phases into activities or tasks they described as sequential. Major phases thus appeared to be quasi-independent, with little flow from one professional team to another. A few exceptions are as follows:

(a) Technicians often receive instructions from a scientist for building a piece of equipment or taking readings, perform the work, and return to the scientist with the results.

(b) Scientists sometimes perform service functions for other scientists in addition to performing their own research. For example, one professional may design electrical circuits for others or serve as a computer programming "trouble-shooter" for others.

(c) Supervisors appear to participate especially heavily in the overall design of the research and in reporting or writing up of results. They appear to be less involved than the other professionals in minute technical details except when special problems arise or the work reaches some particularly critical stage. Thus the work flow can broadly be described as moving from the supervisor to the other researchers and back to the supervisor.

(d) Time dependencies sometime occur when one device depends on the development of another for calibration, or a system development depends on the refinement of a component such as a device or a complex computer program.

(2) Continuous observation of the work process

(a) The method used in continuous observation of the work process consisted of an analyst making notations every two minutes on the activities of a particular scientist. Five analysts acted as observers providing a total of ten man-days of continuous recorded observation. Scientists to be observed were selected from all five programs early in the study so that each program is represented by two days of observation. Approximately 2,400 observations were made by this means. Two factors were recognized as biases in the data: [1] Two days constitute too short a time span in the life of a scientific research program, and are undoubtedly non-representative of the work processes that typify the program over its entire life span. This bias was somewhat mitigated by the observation of five different programs in various stages of development. [2] Continuously observing a person's behavior is likely to cause changes in that behavior. This unavoidable bias was mitigated by observation of people considered least likely to be threatened by the process. Because of their concern about this factor, the program managers in some cases chose to be observed themselves. Hence the data is divided into two categories, supervisory and non-supervisory use of time. The time covered is divided about equally between the two categories, as the summary data in figure 2 show.

The summary of the continuous observations shows non-supervisory personnel spending about 43% of their time setting up experiments and taking readings and 25% of their time talking about the technical work. They spend little time (4.2%) writing. The supervisors, by contrast, spent very little time setting up experiments and taking readings (2.5%) but a good deal of time (24.6%) writing. Like the non-supervisory personnel, they spent a major portion of their time (30.5%) talking about the technical work.

(3) Work Sample

(a) Method

The work sample was conducted for two weeks (February 5 to 16). Science managers, scientists, engineers, and technicians were observed four times on each of the ten days beginning at times which were randomly selected. Four programs (Cryoelectronics, Cryogenic Metrology, Quantum Electronic Frequency Standards, and Laser Parameter Measurements) and 29 people working in these programs were included in the work sample. Starting at randomly selected times, an observer followed a route selected randomly from among several routes between buildings and inside the buildings where the programs were carried out. He then recorded observations on the activities of each person in the sample population. The total sample consisted of forty runs and 1158 observations. An average run time was 54.6 minutes. Statistical significance was not achieved for all of the elements of the sample, but the data collected agreed generally with the data collected in the continuous observation.

| Activity | Supervisory Total Min. | Supervisory Item % | Supervisory Category % | Non supervisory Total Min. | Non supervisory Item % | Non supervisory Category % |
|---|---|---|---|---|---|---|
| 100 Experiment/Calibrate | - | - | 2.5 | - | - | 42.5 |
| 110 Experiment: Set up | 62 | 2.5 | | 761 | 31.8 | |
| 120 Experiment: Take readings | - | - | | 222 | 9.3 | |
| 130 Experiment: Tear down | - | - | | 33 | 1.4 | |
| 200 Reading | - | - | 4.6 | 4 | .2 | 7.0 |
| 210 Reading Technical Materials | 35 | 1.4 | | 61 | 2.6 | |
| 211 Reading for Design purpose | 13 | .5 | | 88 | 3.8 | |
| 212 Reading for analysis purposes | 15 | .6 | | 8 | .3 | |
| 220 Reading for administrative purposes | 50 | 2.1 | | 2 | .1 | |
| 300 Writing | - | - | 24.6 | - | - | 4.2 |
| 310 Writing for Technical purposes | 288 | 11.8 | | 66 | 2.8 | |
| 311 Techical writing: design | 227 | 9.3 | | 8 | .3 | |
| 312 Technical writing: analysis | 43 | 1.8 | | 2 | .1 | |
| 320 Writing for administration | 42 | 1.7 | | 23 | 1.0 | |
| 400 Talking | 8 | .3 | 30.5 | - | - | 25.3 |
| 410 Technical talk <u>in</u> parent agency | 221 | 9.9 | | 381 | 16.0 | |
| 411 Technical talk: design | 122 | 5.0 | | 118 | 5.0 | |
| 412 Technical talk: analysis | 179 | 7.4 | | 41 | 1.8 | |
| 420 Technical talk <u>outside</u> agency | 52 | 2.1 | | - | - | |
| 421 Technical talk outside: design | - | - | | - | - | |
| 422 Technical talk outside: analysis | - | - | | - | - | |
| 430 Administrative talk | 142 | 5.8 | | 60 | 2.5 | |
| 500 Telephone use | - | - | 9.3 | - | - | 3.2 |
| 510 Telephone inside agency | 93 | 3.8 | | 73 | 3.0 | |
| 511 Telephone inside: design | 10 | .4 | | 4 | .2 | |
| 512 Telephone inside: analysis | - | - | | - | - | |
| 520 Telephone outside agency | 123 | 5.1 | | - | - | |
| 521 Telephone outside: design | - | - | | - | - | |
| 522 Telephone outside: analysis | - | - | | - | - | |
| 600 Walking | 62 | 2.5 | 2.5 | 46 | 1.9 | 1.9 |
| 700 Not present | - | - | 12.9 | - | - | 3.3 |
| 710 Not present, in other on-site location | 52 | 2.1 | | 78 | 3.3 | |
| 718 Another area on-site housekeeping | 93 | 3.8 | | - | - | |
| 740 Not present, Prof. development | 84 | 3.5 | | - | - | |
| 750 Non-prof. mtg. (EEO) | 66 | 2.7 | | - | - | |
| 770 Not present - non contributing | 20 | .8 | | - | - | |
| 800 Administrative Housekeeping | 99 | 4.1 | 4.1 | 43 | 1.7 | 1.7 |
| 900 Non-contributing | 231 | 9.0 | 9.0 | 268 | 11.2 | 11.2 |
| | 2432 | 100.0 | 100.0 | 2390 | 100.0 | 100.0 |

Figure 2
Continuous Observations (two days)
12/5 - 12/6/72

(b) Findings

The Work Sampling Summary (figure 3) shows time spent for all categories of personnel in all of the programs: 41.8% in design and conduct of experiments, 12.1% in building experimental equipment 12.1% in analysis and reporting results, 6.7% in consultation, 6.7% in managerial activities, and 4.8% in professional development. About 5% of the total time went into various routines such as calibrations and general housekeeping.

C. Representing the Process

(1) The need for integration and abstraction

As data were analyzed and summarized in the form of tables designed for review and assimilation, the PROMEX team felt the need to integrate the whole body of information into some readily understood form. Also, there was a need for [1] abstraction to what was general in the whole process from input through work process and outputs to impact, and [2] representation or description of the sequence of events and to the degree possible, the causal chain from input to result.

(2) The development of a model

The flowchart or model (see figure 4) for the NBS scientific knowledge and technology delivery system was developed in an attempt to show some of the relationships between inputs, work processes, outputs and impact. Data from the preliminary work sampling survey were used to show how costs were distributed during the work process. Questions which should be addressed in order to measure productivity are: If we put x resources into project y, what is the probability that the value of the impact will be greater than x? How much greater than x? How can one determine whether the impact/input ratio is greater or less than one? The main purpose of the model is to help relate input, output, and impact in terms of cost and value.

Figure 4 displays the over-all productivity process from input - through work process - through output - to impact. It shows, for those activities included in Promex, the distribution of each input dollar of which 45 cents goes into direct personal services. It further shows the distribution of direct technical manhours and dollars across technical activities.

One unanticipated benefit of the model was that it led to an analysis of the outputs into three primary categories: media for output dissemination; content of the output; and by-products of the work process.

3. Inputs to the Process

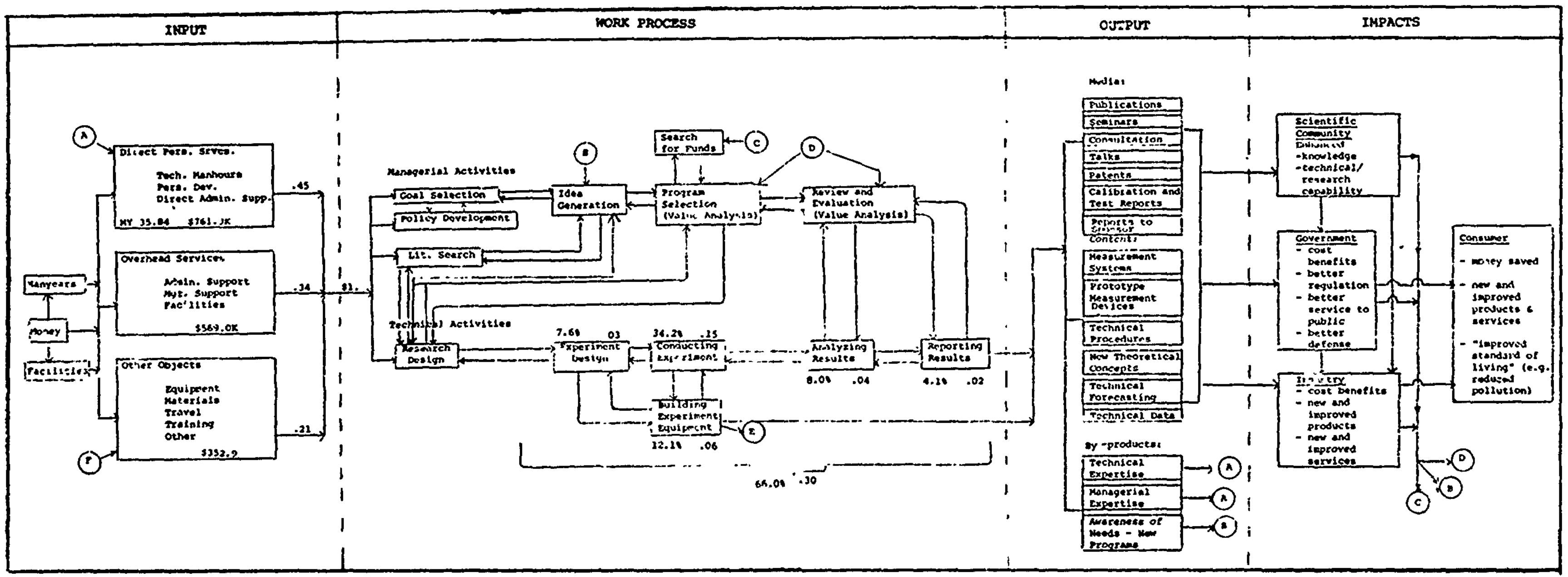
For purposes of simplifying the Promex Model, inputs were grouped into three categories: direct personal services, overhead services; and other objects. These input categories are defined as follows:

| Activity | Mgrs. | Profs. | Techs. | CE | CM | Q-E FS | LPM | Total |
|---|---|---|---|---|---|---|---|---|
| Experiment Design | 4.5 | 9.6 | 2.5 | 10.3 | 6.1 | 2.6 | 9.3 | 7.6 |
| Conduct Experiment | 23.5 | 34.0 | 44.1 | 24.7 | 55.7 | 18.1 | 35.0 | 34.2 |
| Bldg & Main. Exp. Equip. | 1.3 | 9.9 | 29.4 | 16.1 | 3.2 | 25.4 | 6.4 | 12.1 |
| Analyze Results | 7.1 | 9.9 | 1.0 | 4.3 | 6.4 | 18.1 | 8.0 | 8.0 |
| Report Results | 11.4 | 3.6 | --- | 5.3 | 3.2 | .5 | 6.3 | 4.1 |
| Consultation | 13.9 | 7.0 | --- | 6.6 | 1.3 | 2.0 | 17.7 | 6.7 |
| Calibration | --- | .6 | --- | --- | 1.6 | --- | --- | .4 |
| Professional Development | 3.1 | 5.9 | 2.0 | 10.6 | .5 | 2.0 | 3.0 | 4.8 |
| Managerial | 20.8 | 4.8 | 2.0 | 5.5 | 8.3 | 8.8 | 4.2 | 6.7 |
| Administrative | 6.9 | 4.7 | 3.0 | 4.3 | 4.8 | 8.3 | 2.5 | 4.7 |
| Present Activity Unknown | 5.6 | 6.9 | 12.2 | 8.8 | 7.0 | 10.4 | 4.2 | 7.6 |
| Non-Contributory | 1.9 | 3.1 | 3.5 | 3.5 | 1.9 | 3.6 | 3.4 | 3.1 |
| Total | 100.0 | 100.0 | 99.7 | 100.0 | 100.0 | 99.8 | 100.0 | 100.0 |

Figure 3
Work Sampling Summary - % of time spent
2/5/73 - 2/16/73

Figure 4

PROCESS OUTLINE CHART




1. Work Sample Data

| | % | $ |
|---|---|---|
| Technical Manhours | | |
| Experiment Design | 7.6 | .034 |
| Conduct Experiment | 34.2 | .154 |
| Bldg. & Main. Exp. Equip. | 12.1 | .055 |
| Analyze Results | 8.0 | .036 |
| Report Results | 4.1 | .018 |
| | 66.0% | $.297 |

| | % | $ |
|---|---|---|
| Other Direct Manhours | | |
| Consultation | 6.7 | .030 |
| Calibration | .4 | .002 |
| Prof. Dev. | 4.8 | .022 |
| Managerial | 6.7 | .030 |
| Administrative | 4.7 | .021 |
| Present Act. Unknown | 7.6 | .034 |
| Non-Contributory | 3.1 | .014 |
| | 34.0% | $.153 |
| Total | 100.0% | $.450 |

2. NOTE: Cost Data is for FY '73.

Letters in circles indicate outputs that feedback through the process.

° direct personal services -- includes cost for technical manhours, costs for professional development (except for formal training fees), and administrative support costs which are charged directly to technical cost centers.

° overhead services -- includes administrative and managerial labor costs and facilities costs; all are charged to either division, institute, or bureau overhead cost centers.

° other objects -- includes costs for equipment, materials, travel, training, etc., costs which are charged to technical cost centers.

Figure 5 shows estimated man years and total costs for FY '73 and FY '74 by program. (From NBS Project Reports)

| | Man years | | Costs | |
|---|---|---|---|---|
| | '73 | '74 | '73 | '74 |
| Cryoelectronics | 9.85 | 11.9 | 427.5 | 550.5 |
| Cryogenic Metrology | 9.89 | 6.8 | 418.9 | 306.0 |
| Frequency Standards | 9.3 | 7.5 | 383.0 | 324.0 |
| Laser Parameter Meas. | 7.4 | 12.4 | 450.1 | 580.1 |

Figure 5
Estimated Man Years and Costs FY '73 and FY '74

Figure 6 shows actual costs for labor, overhead, and other objects for FY '73. (From NBS accounting reports).

| | Labor | Bureau O/H | Inst. O/H | Div. O/H | O/O | Total |
|---|---|---|---|---|---|---|
| Cryoelectronics | 217.2 | 80.4 | 7.4 | 82.7 | 42.7 | 430.4 |
| Cryogenic Metrology | 182.2 | 67.4 | 6.1 | 69.4 | 93.8 | 418.9 |
| Frequency Standards | 181.8 | 54.7 | 6.1 | 69.5 | 70.3 | 382.4 |
| Laser Parameter Meas. | 180.7 | 50.9 | 5.9 | 68.5 | 146.1 | 452.1 |

Figure 6
Actual Costs FY '73

After examining the work process and developing the flow model, it appeared quite feasible to satisfy some of the measurement factors as defined in <u>Dimensions of Productivity</u> (figure 1, page 7). The value of input could clearly be quantified. Efficiency in generation of output could be expressed in terms of use of time. Under 'value of output' the volume and even the quality of output might be specified, but the utility of output remained elusive. One of the techniques scheduled for use at the time the "blueprint" was published was value analysis. Value analysis appeared to offer one approach to quantifying utility of output (impact).

## 4. Value Analysis

Value analysis is a technique for quantifying anticipated benefits, advantages, and disadvantages of alternative courses of action. It is particularly useful in focusing on specific factors or concerns in the decision-making process. The method usually involves the development of specific criteria and then reviewing each alternative according to the criteria to obtain an overall value (score) for each alternative. The scoring process usually requires the development of weights to compensate for differences in the relative importance of the criteria.

If consensus on criteria and weights is achieved, it is possible to use this technique in decision-making relative to:

- program selection,
- program evaluation,
- reprogramming decisions.

To explore the idea of using value analysis within the NBS managerial context a task force was established in Boulder under the direction of the Deputy Director, IBS/Boulder. The task force consisted of all Boulder division chiefs plus 2 staff assistants. The following outline details the objectives established for the task force:

(1) Develop a set of criteria for program selection and evaluation in terms of measurement of accomplishment.

(2) Select methods for applying criteria.

(3) Develop weights to be assigned to criteria in experimental value analysis matrix.

(4) Apply criteria, testing different application methods, to programs included in the experiment.

(5) Analyze reasons for discrepancies between expectations and current accomplishments.

(6) Evaluate usefulness of various methods for applying criteria to program selection and evaluation.

(For a more detailed outline of Task Force Objectives, see Appendix H, "Value Analysis").

The task force developed a set of criteria which were tested during the program reviews in Boulder. (See Appendex E, "Statistical Analysis of Rating Results," and Appendix H, "Value Analysis.") In the actual application of the value analysis technique the following criteria were used:

(1) Relevance to National Needs

(a) Primary legislative responsibility: Does the program promote national capability for physical measurements to the accuracy or precision needed?

(b) National goals and needs of society: Are the significance and urgency of the problems addressed great in terms of their impact on the nation?

(c) Payoff: Are the anticipated outputs significantly greater than anticipated inputs?

(d) Leverage: Can NBS have a unique and substantial impact on the problem? Who is waiting for the results? What will they mean to him?

(2) Institutional Health

(a) Probability of success: Is the problem well analyzed? Does past performance point toward success? To what extent is the field ripe for exploration?

(b) Resources: Do the funds, leadership, and technical capabilities exist to support such a research effort?

(c) Technical merit: What is the technical quality of the output? Does the program enhance NBS stature? Does the program draw from or contribute to other fields?

(d) Staff welfare: Is there opportunity for desirable individual growth? Development of new skills? Opportunity for scientific contribution?

Figure 7 is a copy of the actual rating form which was used.

At the Base Program Review held in Boulder the criteria developed by the Value Analysis Group was used to evaluate the Boulder programs. Seven science managers from management levels above the presenters, completed the rating sheets. The data from the evaluation exercise were used to test these hypotheses:

° Weighting the criteria will significantly effect the ranking of the programs.

Program ____________________

Presenter ____________________

Rater ____________________

| | None | Low | Medium | High | Very High |
|---|---|---|---|---|---|
| **1. RELEVANCE TO NATIONAL NEEDS** | | | | | |
| a. Primary legislative responsibility: Does the program promote national capability for physical measurements to the accuracy or precision needed? | ☐ | ☐ | ☐ | ☐ | ☐ |
| b. National goals and needs of society: Are the significance and urgency of the problems addressed great in terms of their impact on the nation? | ☐ | ☐ | ☐ | ☐ | ☐ |
| c. Payoff: Are the anticipated outputs significantly greater than anticipated inputs? | ☐ | ☐ | ☐ | ☐ | ☐ |
| d. Leverage: Can NBS have a unique and substantial impact on the problem? Who is waiting for the results? What will they mean to him? | ☐ | ☐ | ☐ | ☐ | ☐ |
| **2. INSTITUTIONAL HEALTH** | | | | | |
| a. Probability of success: Is the problem well analyzed? Does past performance point toward success? To what extent is the field ripe for exploration? | ☐ | ☐ | ☐ | ☐ | ☐ |
| b. Resources: Do the funds, leadership, and technical capabilities exist to support such a research effort? | ☐ | ☐ | ☐ | ☐ | ☐ |
| c. Technical merit: What is the technical quality of the output? Does the program enhance NBS stature? Does the program draw from or contribute to other fields? | ☐ | ☐ | ☐ | ☐ | ☐ |
| d. Staff welfare: Is there opportunity for desirable individual growth? Development of new skills? Opportunity for scientific contribution? | ☐ | ☐ | ☐ | ☐ | ☐ |
| 3. Rate the quality of the presentation: | ☐ | ☐ | ☐ | ☐ | ☐ |
| 4. Rate the extent to which the quality of the presentation may have effected your evaluation: | ☐ | ☐ | ☐ | ☐ | ☐ |
| 5. Rate your technical knowledge of the field or program area covered in the presentation: | ☐ | ☐ | ☐ | ☐ | ☐ |

Figure 7
Criteria for Evaluation: Value Analysis

° Division chiefs will not bias their ratings in support of their own programs.

° Programs which emphasize relevance to national needs criteria will grade out higher than programs which emphasize institutional health criteria.

° Quality of presentation will significantly affect the ranking of the programs.

Using data generated at the Base Program Review we reached the following conclusions relative to the above hypotheses:

° Weighting the criteria did not significantly affect the ranking of the programs.

° Division chiefs did not bias their ratings in support of their own programs.

° In general programs which emphasize national need criteria did tend to grade out higher than programs which emphasize institutional health criteria.

° Quality of presentation tended to affect the raters only slightly.

The main use of the rating sheet was to develop a priority list of rated programs. We then compared this data with the intuitive ranking of programs which was obtained separately from the raters. With one major exception there were no differences in quartile placements. Different weights and different scoring systems had little effect on the final list of programs.

Major findings are: 1) this kind of rating scheme can be used in the selection and evaluation process; 2) different sets of criteria are not required for selection and evaluation; 3) different sets of criteria are not required for different programs; e.g. calibration work vs. less applied work. Correlations on the effect of presentation style and correlations on the effect of the rater's own knowledge and background indicated no major biases from these factors. Tests for exclusiveness and comprehensiveness have shown the criteria to be more than adequate. (See Appendix E, Statistical Analysis of Rating Results.)

5. Outputs of the Process

The initial interviews yielded the following list of program outputs: 1) publications, 2) devices, 3) calibrations and tests, 4) capability to provide calibrations and tests, 5) improved measurements of physical quantities or fundamental constants, 6) new performance or safety standards, 7) new or improved services, 8) computer programs. Discussion of "what is an output" with many NBS scientists and engineers generally led to the conclusion that most devices, improvements, new standards, etc., are described in publications. The consensus seemed to be that publications, either in themselves or what they describe, represent about 80% of the primary output of NBS.

Most of the earlier output analysis effort focused on the publication process. Four approaches to the evaluation of publications in a quantitative and/or qualitative fashion were explored: publication classification, citation indexes, reprint requests, and journal classification.

Publication Classification - This method would classify publications according to four factors:

1) publications ranked according to content type (theory, synthesis of current knowledge, descriptions of experiments, etc.;

2) ranked according to types of literature (scientific journals, trade journals, etc.);

3) type of scientific literature in which each type of publication would have its greatest impact;

4) the frequency of printing of each type of publication.

As a result of very negative reaction by members of the scientific staff, the publication classification measure was dropped. The main argument against the method was that it was overly biased in the direction of professional and personal prejudices.

Citation Indexes - The citation indexes are designed to provide a measure of the impact of a journal article or other publication. The index shows the number of times per year an article is cited. The impact of the article is considered to be directly related to the number of times it is cited. Citation data can be purchased from the Institute for Scientific Information (ISI). The convenience and availability of the citation indexes make . e advantageous. However, as a sole measure of publication impact, citation indexes are inadequate. There are several problems associated with the manner in which the indexes are compiled and with the assumption that impact is directly related to number of citations. For example, citation indexes give no indication of how an article is used; an article whose analysis is wrong may be widely cited by those trying to correct it.

Reprint Requests - Reprint requests also measure publication impact. Several NBS staff members feel that there is a strong relationship between reprint requests and the impact of an article. Data on reprint requests can be collected easily and cheaply.

Journal Classification - This method is designed to evaluate publication quality. Rankings for journals, according to quality, were derived by use of a questionnaire presented to the scientific staff. By inference, these journal quality rankings give an indication of the quality of articles published in them.

Data were collected and compared for a sample of one hundred articles using the reprint request, citation index, and journal classification methods. Citations and reprint requests data for four ranges in the sample showed a direct correlation. Data on journal type and rank showed little direct relationship to the other two measures. (For more details on these issues see Appendix F, "Output Analysis: Publications").

## 6. Impact of the Process

Impact is achieved through a wide variety of activities including dissemination of research results, provision of services, demonstration or dissemination of devices, publication of standards, personal contact with research workers in other laboratories, participation in conferences, committee work, etc.

Relatively little quantitative data on impact was expected to be generated by PROMEX. More comprehensive data were anticipated from the National Measurement System (NMS) studies conducted independently by the Institute for Basic Standards (IBS). These studies included surveys of the organizations and individuals who use IBS outputs to evaluate the value of such outputs. Data from the NMS studies might then be coupled to knowledge gained through PROMEX to describe the value of NBS programs.

The NMS studies have the following objectives:

- Demonstrate structural organization and interface with "users";
- Show current distribution of effort;
- State measurement responsibilities delegated to division;
- Delineate importance on technologies (Quantify with economic data and show social needs);
- Delineate importance of measurement needs;
- Evaluate planning of current programs; and
- Predict future program directions and changes.

## 7. Communications Study

Gathering information from the outside is an essential part of the planning process, just as disseminating information is an essential output. Inside communications are a fundamental part of the work processes in between planning and accomplishment.

Because the flow of information in a research effort has a major influence on both effectiveness and efficiency, an analysis of communications was included as a segment of the productivity study effort.

The Study Procedure. The first step in the study was to develop a questionnaire to measure the perceptions of the staff about the work situations. Previous research studies were reviewed to find validated questions which could be used in this study. Questions were selected from the communication climate study by Litwin (1968), a job characteristic study by Hackman and Lawler (1971), and a study by Friedlander (1969). Bipolar adjective word pairs were selected from Berlo, Lemert, and Mertz (1970). In addition, questions were developed to measure perceptions concerning the barriers to communication.

The questionnaire was field tested in the Management and Organization Division. After revisions, the questionnaire was administered to two groups at Gaithersburg -- nine participants in one group and sixteen in the other one. The Boulder groups then completed the questionnaire.

Data Analysis. Analysis of the resulting data concentrated on six key dimensions: Source Credibility, External Feedback, Autonomy Task Identity, Variety, and Internal Feedback.

The Multidimensional Scaling Technique. A technique which could be used to analyze the interrelationship of perception data was sought for this study. A relatively new analysis technique, multidimensional scaling, was found to satisfy the requirements. The technique not only allows this type of analysis but also allows for graphic display of the results. A number of versions of this technique are presently in operation. One version was tested and several others were studied before the MDPREF (Multidimensional Preferance) version was selected.

MDPREF is a vector model of multidimensional scaling. This computer program finds the interrelationship of the stimuli (question/bipolar word perception scores) and defines them in space. A plane which best fits these stimulus points in two dimensions is then found and vectors for each participant are then "fitted and drawn" through these points. Each vector results from projections of the stimulus points upon the vector line and represents the "best rank ordering of these points for a particular participant. (See Appendix D, "Communications Study")

In this study, stimulus questions were used to measure the characteristics of the work situation and bipolar adjective pairs were used to measure the credibility characteristics of supervisors. A questionnaire was administered by a member of the PROMEX working group. (See Appendix D) Perceptional judgments were made by the members of each program unit.

## General Conclusions

The following are capsule summarizations of the data generated from the communications questionnaire.

### The Supervisor Profile --

The majority perceived their supervisor as really knowing his job, as being a delegator of responsibility, and as being trustworthy. They also perceived a lack of personal growth opportunities for themselves and a lack of clear assignments, feedback, and pressure to produce from their supervisors.

### Communication Down --

The majority perceived a mutual trust between themselves and their supervisor, their supervisor as frank and open, as committed to their programs, and as giving adequate responses to questions.

They also perceived that their opinions are not often asked on decisions already made and if they were, the decisions were not modified. They perceived policies and procedures as difficult to understand and lacked trust in both the Bureau and Institute level of management. To a lesser degree, they also perceived a lack of upper management commitment, a lack of clear assignments and a loss of non-technical communication as information filters down the organizational hierarchy.

### Communication Up --

The majority perceived a mutual trust between themselves and their supervisors whom they perceive as frank and open, as giving prompt answers, and as easy to ask for help. To a lesser degree, they also saw their supervisors as committed to their programs and they perceived opportunities to give additional information to their supervisors.

They also perceived less trust in Bureau and Institute management and saw infrequent opportunities to present their case to higher levels of management. To a lesser degree, they also saw a lack of upper management commitment.

### Horizontal Communication --

The majority perceived a good working relationship with their co-workers and saw themselves as members of a well functioning team. To a lesser degree, they saw trust within their group and they did not feel their work suffered from lack of communication with their peers.

They did not perceive rewards for communicating with other units. To a lesser degree, they also did not see much opportunity to have many work discussions with other units.

Independent/Dependent --

In general, both of these groupings perceived a high satisfaction with their supervisor and their co-workers. They also both perceived less encouragement to take calculated risks and less pressure to improve.

The independent group, however, saw relatively more satisfaction with their supervisor, more autonomy, more task identity, more responsibility and more of an exhibition of trust from their supervisor. They also saw relatively less discussion with other units, less pressure to improve or produce, less peer recognition and less discussion with their peers.

The dependent group saw relatively more satisfaction with their co-workers and a better working relationship with their peers. They also saw relatively less emphasis on solving problems on their own.

Characteristics important to the work outputs --

The majority perceived commitment by their supervisor, good working conditions, challenging goals and assignments and that their outputs were important.

They also saw a lack of positive impact of their outputs on society, a lack of good equipment, and a lack of understanding of organizational policies and procedures.

Central Communication Conclusions

The general feeling about communications was that it is critical, poor, often the cause of error and low efficiency, and takes up between 15% and 80% of staff members' time. Most of the program managers indicated that a higher level of management than theirs made decisions as to which programs would be funded and pursued and which would not.

One of the primary problems related to downward communication is the tendency for organizational members at different levels of the organization to perceive the same information differently (e.g. individuals at a higher level perceive giving instructions and those at a lower level perceive them as information or advice). The questionnaire data reveal a problem in this area for the Promex programs. Many technical workers do not understand Bureau policies and procedures, work assignments are often unclear, and feedback is generally absent.

The primary difficulties with upward communication are often associated with the mobility aspirations of the communicator and his perceptions of the person to whom he is communicating. An ambitious communicator may tend to withhold or distort messages so as to protect and enhance his relationships with higher level individuals. This tendency is reduced to the degree of trust and confidence he perceives between himself and the receiver. The

Boulder group, as a whole, saw being promoted as relatively unimportant. In addition, they perceived a high degree of mutual trust between themselves and their immediate supervisors. These conditions would indicate a good climate for upward communication at the section or team level. However, the fact that the majority perceived relatively less trust in their Bureau and Institute management, saw a relatively lesser opportunity to present their case to higher management, and hesitate to volunteer negative information to any level, indicates an upward communication problem.

There is evidence that specialization of functions tends to impair horizontal communication and that interpersonal sensitivity does not significantly improve it. The big question in horizontal communication is "what's in it for me?" The reward of the organization is the key factor. The majority of the Boulder group saw relatively little reward for work related communication with other units. To a lesser degree they also saw themselves with relatively few opportunities for discussions with other units. Although they generally agreed, at a moderate level, that their work did not suffer because of lack of communication with their own peers or those of other units; this is an area with potential for improving productivity. For example, on projects where a cooperative effort between units is desired, needed horizontal communication could be enhanced by "building in" positive rewards (such as joint recognition) into the project plan.

## III. RESULTS

### 1. Examination of findings

In each of the program areas, when first phase interviews had been completed, individual work assignments were identified. For the most part, work assignments are made as entities to individuals and correspond to major parallel phases of the program. Work does not normally flow sequentially from individual to individual. Within major phases we did not have complete descriptions of work steps, e.g., for "build laser" or "design heat exchanger," and needed to gather data through further interviews to complete the descriptions. In addition, we needed to gather information on the time required for administrative tasks (personnel, budget, and equipment) and off-site activities (training, seminars, and field trips). Because of the diversity of information sought, the level of detail required, and the desire to maintain minimum interference with technical projects, we attempted to obtain much of this information by work sampling. Using this technique, we were able to develop information on interruptions to the work process, and on distribution of effort within the work process.

While many obstacles to the efficiency of operations were cited, the interviewees were not generally concerned with efficiency except in cases where the lack of it would be so extreme as to impair effectiveness. The main thrust of their efforts is getting a high quality job done. The question of what events or outputs might indicate effectiveness was answered in terms of the use of outputs.

The list of efficiency impediments in Appendix 3 covers certain areas where further study might help improve the efficiency of Bureau services. We referred the points raised by the interviewees to those responsible for providing such services. Otherwise, we dropped efficiency factors as a consideration of the experiment.

The information gathered during the interviews established that formal work plans could be developed for all the programs included in the experiment. In some cases such plans had already been developed. The actual work plans that were developed from the interview material differed with respect to the amount of detail, the degree of activity definition, and precision of estimated completion dates. While the exercise of producing these plans from the interview material has not produced work plans which can be used in their present form, with some additional effort these plans could be extended to include more definite milestones and estimated costs and completion dates. The question of whether it is worth the effort to develop detailed work plans can best be answered by developing such plans and monitoring progress against them. The experience gained in this process would then lead to recommendations on the usefulness of formal planning in research programs of this type.

An attempt has been made to derive descriptions of outputs and impact from the interview data. For purposes of the productivity study, we needed to clarify definitions of outputs and identify outputs which would be quantified for comparison with inputs. Outputs are the subject of further analysis.

In the tentative list of specific outputs in Appendix G, an output is considered to be any of the following:

(1) a publication or other dissemination of research results;

(2) a completed instrument or device;

(3) calibrations and tests or the capability to perform them;

(4) completion of improved measurements of physical quantities and/or fundamental constants;

(5) new standards of performance and safety;

(6) establishment of new services, such as the Measurement Assurance Program, and others;

(7) reducing the uncertainty in a measurement process;

(8) computer programs for dissemination

or any other tangible results of the Bureau's work. New information obtained from research is considered an output when it is disseminated in some way outside the immediate group working on the research. Also, outputs may be devices or techniques which are by-products of the experimental work, as well as those which are end objects of the work.

The objective of most NBS programs is to meet a need in the scientific community, government, industry, or society. This is to say that NBS' programs should result in a benefit or saving or other desirable impact in the particular area of need. By measuring output we are focusing on media, an intermediate state or result, which may or may not be related to real impact. (See figure 4, Process Chart)

It could be argued that while output is not as meaningful an attribute in measuring program productivity as impact, because output is easier to measure and quantify, it should be used as a substitute for impact. In many cases if the measurement of a particular attribute is very difficult, the use of a surrogate measure may be desirable. For example, an economist really interested in measuring economic well being in a population may use the surrogate measure of gross family income.

However, the impact of NBS programs has little relationship to the countable output because there are instances where impact occurs without output, where impact precedes output by months or years, and where output occurs without achieving impact. Impact occurs when valuable information is transferred to a person who uses it. This lack of correlation between countable output and impact exists partly because there is a well developed informal communication network between NBS scientists and the outside scientific community, particularly among those engaged in closely related areas of research. Consequently, considerable amounts of output resulting in further, valuable information is often communicated among those doing related research and can have impact long before countable output occurs. For example, with delays in editing, proofing, and publication the information may be disseminated informally a year or more before the formal, measurable output (a publication) occurs.

The importance of informal communication media in disseminating scientific information which has significant impact is discussed in an article by Raymond Isenson entitled "Technological Forecasting Lessons from Project Hindsight."[1] The Isenson study which traced the idea flow for a number of high impact inventions and innovations concluded that:

> *The interesting factor is the relative dominance of informal person-to-person confrontation as a propagating mode. Surprisingly, even in the case of science, the informal link is about as important as the published paper. This observation suggests that the growth of technology is sensitive to the relative ability of the involved scientist to communicate freely. It also suggests that technical publications are not adequate to provide the requisite degree of communciation.*

Data from the Isenson study indicate the following:

| Activity | Personal Contact | Publication of Report | Seminar or Symposium |
|---|---|---|---|
| Science | 45% | 53% | 2% |
| Technology | 64% | 33% | 3% |
| Engineering (design) | 79% | 21% | 0 |
| Engineering (mfg. technique) | 77% | 23% | 0 |

At the present time informal media are not tested as countable output. However, even if informal output were included in the reporting system there are other problems associated with using output as a productivity measure.

First, it is entirely possible to have countable output when impact is negligible or non-existent, e.g. many articles or talks which review or summarize research but present nothing novel, outputs which are

[1]Cited in Technological Forecasting for Industry and Government Methods and Application, James R. Bright (ed.), p. 49.

addressed to particularly esoteric topics, and outputs addressed to audiences where impact or penetration is extremely difficult. Some unpredictable factors affecting impact are variation in the quality of output, appropriateness of output to the audience addressed, receptivity of the audience, and type of output media.

Although no statistical studies have been done, in PROMEX, the following arguments indicate a poor correlation between output and impact.

(1) There is little relationship between the volume of outputs and the value of impact. For example, one report to one sponsor (and distributed no where else) can have significant impact on the sponsor and/or sponsoring industry, whereas, a popularly written review or summary article having wide circulation and many reprints, may result in little impact. Further, more feedback often occurs on popularly written articles which simply draw conclusions or treat a subject in a general way, than on well done scientifically substantiated articles which presumably have impact in the scientific community.

(2) If output is used to measure the productivity of programs or individuals, programs will gear themselves to produce appropriately. Conversely, if productivity is measured using impact, programs will be motivated in that direction.

(3) Inputs, measured in dollars, are incompatible with unit output measurements. (It is not our goal merely to produce outputs as cheaply as possible.) Impacts, however, can sometimes be measured in delayed dollar values, and therefore, can be compatible with input measurement.

In ongoing research or the planning of research, there are varying degrees of information for use in anticipating impact. (For a list of impacts, see Appendix G.)

For discussion of findings from the communications study, see pages 24-26.

## 2. Conclusions

Past attempts to measure R&D have for the most part focussed on statistical units of measure, and particularly on various aspects of publications. Those who have made such attempts, whether through direct counting, citation indices, or weighting the quality of journals, have generally concluded that each attempt brings them closer to a sophistication in counting and weighting which will eventually enable them to use these units of output as a reliable indicator. However, what they are trying to measure has never been made sufficiently clear.

In this experiment in measuring productivity at NBS, in order to help clarify the 'what' part of the measurement problem, a model (figure 4) of the R & D process was constructed, built around a group of rather "researchy " applied programs. Our concern was primarily with the relationship between the cost of input (or resources) and the values of outputs.

Outputs themselves presented a problem of definition. Some outputs were inputs to another step in the process though such outputs were usually described in publications or other output media.

Since we were neither concerned with an output's contribution to a profit unit nor its value in the marketplace, we had to look closer in order to measure achievement of objectives, to the impact of our outputs on various outside sectors.

The distinction between media descriptions of work as outputs and the actual content of work became necessary because what we wanted to measure was the value of work in terms of impact. This impact often could not be traced back through media descriptions (e.g., publications, invited talks, conferences, etc.). There was no necessary connection between appearance in the media and the impact of the work.

The use of media counting, and particularly the publication record, has been used as a standard measure to indicate the professional status of individuals and the technical health of organizations. The publication record invariably accompanies applications for employment or requests for promotion. In discussing the performance of a technical group the number of publications is always cited as an indicator of the relative technical health of the group. It is a natural step to use the same evidence which substantiated the performance of the group to indicate the productivity of the group since there is an apparent connection between performance and productivity. The premise has been that if sufficient performance data were marshalled, it would somehow comprise a measure of productivity. However, performance is not a sufficient condition for productivity.

The confusion between performance and productivity is readily understandable since most of us think of productivity as the prime factor in performance. For example, the Committee on Federal Laboratories (COFL) report on "Performance Measures for Research and Development," defines performance so broadly that such terms as "productivity, efficiency, impact, and effectiveness". . . are subsumed under the term 'performance'. If one defines productivity in terms of the contribution to the achievement of objectives, then it is clear that within an organization high performers, who produce what they or management have prescribed, may be non-productive in terms of their contribution towards achieving organizational objectives. This phenomenon, which results from misdirected efforts, is especially apt to occur in R&D situations where decisions on what to do are naturally decentralized. Another factor which may affect productivity adversely is the individual researcher's personal objective of peer recognition, which is a strong motivator in career development.

In studying our model of the R&D process we discovered that in attempting to measure productivity, we were not interested in evaluating outputs as indicators of performance levels nor in evaluating those outputs (by-products in the model) which add value to our capital resources. What is critical for us to be able to measure, in order to develop indices of productivity, is the value of the impact which outputs have on various sectors of society. Here we are concerned with the media content regardless of where it appears. Further, we are occasionally concerned with the content of outputs which never appear in the formal, controllable media. Several examples of the kind of problem one finds in identifying the links or transfer mechanisms between the output content and impact occurred during the course of our experiment (Appendix G). In the 10 case studies presented in the TRACES/NSF report[2] the chart on page 4-7 shows informal transfer of knowledge as an important factor in nine cases. This supports the hypothesis that information is informally disseminated and impact occurs before the research is formally disseminated; that is, knowledge and materials (content) are transmitted between organizations by personal contact.

Thus, it is clearly of very limited use to focus attention on approaches to measurement relying on statistical counts of media. A primary requirement for any measurement unit is that it have a necessary connection to what is being measured. There is no necessary connection between counts of media output and research impact. No matter how sophisticated the analysis and synthesis processes become, media units will therefore not serve as reliable units of measure. There simply are no output media units which will satisfy the requirements for useful units of cost/value.

From our experience with PROMEX, two possible techniques may provide useful indicators of productivity. One is a value analysis of criteria for selection and evaluation of results. The other is a survey technique which gets to the specific users of outputs and gets from them experience based estimates of impact.

### 3. Recommendations

The study team has developed, based on findings and conclusions of this study, four primary recommendations for the Bureau:

(1) that value analysis workshops be held for key science managers to evaluate the value analysis experiment looking to adoption of a formal, open system for rating and ranking initiatives and on-going programs.

- The value analysis technique would be adopted as the formal Bureau method for evaluating programs: it would be included as a part of the annual cycle of program planning and evaluation. The system tested in the workshops would be that described in pages 17-20 of this report. The desirability of an open system, one in which ranking of all NBS programs is made known to the staff, is that it will increase understanding of what is expected and support the trust level between Institute and Bureau management and those who work at the Section level. Managers at all

---

[2] U.S. National Science Foundation, Technology in Retrospect and Critical Events in Science, December 15, 1968, pp 4-7.

levels responsible for specific activities would know the criteria against which they would be reviewed and would receive feedback on how their activities are scored and ranked at the Executive Board level.

(2) that further investigation be conducted on the use of two techniques which may provide useful productivity indicators:

- value analysis based on criteria for selection of programs and evaluation of results

- a survey technique which focuses on specific uses of output.

(3) that a study be conducted at NBS of the behavioral and managerial factors involved in the R & D process to assess the impact of these factors on productivity. Areas of special emphasis should include communications, managerial style, organizational climate, and executive and managerial development. If recommendation #1 is accepted, the impact of the open system for rating and ranking programs should also be reviewed. The study would be coordinated by the Management and Organization Division.

The communications study has provided indicators of behavioral and managerial concerns which could be pursued fruitfully. For example, the communications study data reveals that, for the most part, the researchers tested perceive a lack of: growth opportunities, clear assignments, feedback, and pressure to produce. Researchers tested also see themselves as not involved in the decision-making process. Clearly, if widespread, these are serious problems and should be investigated.

Many other organizations, including research oriented organizations, have dealt successfully with such problems. Tools and mechanisms exist which can be applied in resolving the types of problems cited above. The proposed investigation would center around finding those mechanisms which fit NBS needs.

(4) that further work process studies be conducted in other selected areas of the Bureau;

- The programs included in the study represent only a segment of one Institute. Because they were selected with a bias to more "researchy" activities, they are not necessarily typical of work one would find in other institutes, particularly IAT and ICST. The model of the process (figure 4) may not fit a significant percent of the Bureau's activities, therefore the conclusions drawn may be invalid when applied to the Bureau as a whole. Further work should be done to establish similarities and differences between the work involved in the experiment and more applied activities.

# ANNOTATED BIBLIOGRAPHY

Augood, Derek R., *A Review of R&D Evaluation Methods.* IEEE Transactions on Engineering Management, EM-20 (November 1973): 114-120.

Relates to output, impact, and value analysis.

Methods of selecting and evaluating R&D projects involve checklists ranging from a simple list of a few elements to a large number of weighted attributes showing multi-dimensional patterns. A number of indexes of use in analyzing program evaluation data such as the benefit/cost ratio are discussed. The article concludes that a good evaluation method might be achieved by combining impact and risk analysis and incorporating R&D costs and probabilities.

Baker, N.R. and W.H. Pound, *R&D Project Selection: Where We Stand.* IEEE Transactions on Engineering Management, EM-11 (December 1964) 124-134.

Relates to input, output, and value analysis.

An excellent general discussion and review of literature on R&D project selection. Three representative models of project selection are described in detail. Ten different methods of project selection which appear in the literature are summarized. The article notes the extent to which each of these models has been tested. The lack of testing to determine the feasibility and shortcomings of most of the models explains why these methods have not been used.

Bright, James R., ed., Technological Forecasting for Industry and Government. Methods and Applications. Englewood Cliffs, N.J., Prentice-Hall 1968.

Relates to output and impact.

The article by Raymond S. Isenson, "Technological Forecasting Lessons from Project Hindsight" contains useful information on the importance of informal communication media in research and development. Emphasis is given to the need for increased interpersonal and interdisciplinary communication to accelerate technological change. Part of the article focuses on the importance of personal contact as an idea transfer mechanism in science, technology, and engineering.

Burgess, John S., *The Evaluation of a Government-Sponsored Research and Development Program.* IEEE Transactions on Engineering Management, EM-13 (June 1966) 84-90.

Different yardsticks must be employed in evaluating the research and development programs of government laboratories than for industrial laboratories, the paper contends. Factors which should be considered when evaluating the program of a government laboratory are discussed.

Byatt, I.R.C. and A.V. Cohen, An Attempt to Quantify the Economic Benefits of Scientific Research. Science Policy Series 4, London: H.M.S.O., 1969.

Relates to output and impact.

One method by which the economic benefits of fundamental scientific research might be assessed is discussed. The authors present various benefits resulting from scientific research and discuss a method for attempting to quantify the benefits.

Cetron, M.J. and J.D. Goldhar, *The Science of Managing Organized Technology*. IEEE Transactions on Engineering Management, EM-17 (February 1970): 20-43.

The preface and bibliographic reference parts of a forthcoming book, entitled The Science of Managing Organized Technology, by Cetron and Goldhar, are contained in this article. An excellent annotated bibliography covering 25 books dealing with organization structure, management of scientific talent, and creativity is included.

Cetron, M.J. and J.D. Goldhar, The Science of Managing Organized Technology. New York: Gordon and Breach, Science Publishers, 1970.

*When to Terminate a Research Project*, by C.K. Buell, 941-948.

Relates to input, process, and output.

After discussing some of the problems inherent in any attempt to establish guidelines for research and development projects, approximately 20 "warning signs" that may presage project termination are presented. The paper is addressed primarily to industrial R&D, but many of the criteria are applicable to government research.

*Research Budgeting and Project Selection*, by B.V. Dean and S.S. Sengopta, 913-940.

Data obtained from three major chemical companies were used to determine whether scientific analysis could be employed to derive quantitative and objective decision-making procedures in areas where subjective and intuitive means have been used. A model is presented for the general R&D budgeting process. The problem of budget allocation to specific projects is briefly discussed in the paper.

Cetron, M.J., Joseph Martino, and Lewis Roepcke, *The Selection of R&D Program Content -- Survey of Quantitative Methods*. IEEE Transactions on Engineering Management, EM-14 (March 1967): 4-13.

Relates to input and output.

Thirty methods for evaluating and selecting R&D projects are briefly summarized in this excellent article. The attributes,

ease of use, and scientific area of applicability of the various methods are compared and contrasted with each other. An extensive bibliography is included.

Cetron, Marvin J., et al., Technical Resource Management Quantitative Methods. Cambridge, Mass.: The MIT Press, 1969.

Relates to input and process.

A broad view of important quantitative resource allocation techniques now used in government and industry. The systems that are discussed would require modification to fit the situation of any particular laboratory. Some of the most pertinent parts of the book are Chapter 4, Task Selection in Exploratory Development; Chapter 5, Building a Laboratory-Wide Allocation System; Appendix A, Survey of Quantitative Methods; and Appendix B, "Braille," A Case Study on Quantitative Approaches to Aid in R&D Programming.

Committee on Federal Laboratories, Federal Council for Science and Technology. Performance Measures for Research and Development. Two volumes. May 1973.

Relates to output and impact.

This report analyzes the characteristics of Federal government research and development efforts, evaluates the methods of performance evaluation, and draws some general conclusions about the feasibility of measuring R&D performance. Three methods for evaluating performance discussed are the use of publications and activities, peer evaluation, and cost/benefit analysis. A finding of the study is that no general conclusions about the feasibility of measuring R&D performance can be drawn.

Appendix B provides review and critique of publications, citations, and other similar information outputs as a measure of R&D performance. An excellent annotated bibliography about performance evaluation of research and development is found in Appendix D.

Conference on Research Program Effectiveness. Research Program Effectiveness. Edited by M.C. Yovits, et al. Proceedings of the Conference sponsored by the Office of Naval Research, Washington, D.C., July 27-29, 1966. New York: Gordon and Breach, Science Publishers, Inc., 1966.

Relates to input, process, output, impact, and value analysis.

Some parts of this book that are especially pertinent are as follows:

Chapter II, Project Selection in Industrial R&D: Problems and Decision Processes by R.G. Brandenburg;
Chapter III, State-of-the-Art Projection and Long-Range Planning of Applied Research by F. Pardee;
Chapter IV, Program Planning in a Science Based Service Organization by Walter A. Hahn and Harlan D. Pickering;
Chapter VI, Information Systems for the Test of Hypotheses Pertaining to the Theory of Research Management; by Noah S. Prywes and Milton Silver;
Chapter VII, New Tools for Improving and Evaluating the Effectiveness of Research by Sher and Garfield;
Chapter VIII, Evaluating Two Aspects of Quality in Research Effectiveness by Howard M. Vollmer;
Chapter X, A Graph Oriented Model for Research Management by O. Magenster, R.W. Shephard, and H. Grabonski;
Chapter XII, Stochastic Networks in Research Planning by Burton V. Dean;
Chapter XV, Proposal Generation and Evaluation Methods in Research and Exploratory Development by Bernard Sobin and Arnold Proschan;
Chapter XVI, Sources of Ideas and their Effectiveness in Parallel R&D Projects by Thomas J. Allen;
Chapter XVII, Diffusion of Innovations Resulting from Research: Implications for Research Programs Management by A. Shapero;
Chapter XIX, Some Common Concepts and Tentative Findings from a Ten-Project Program of Research on R&D Management by Albert H. Rubenstein;
Chapter XX, The Role of the Research Administrator by C.W. Churchman, C.E. Keuylbosch, and P. Rotoosh;
Chapter XXI, Organizational Factors in Project Performance by Donald G. Marquis and David M. Straight, Jr.;
Chapter XXIII, Conflict and Performance in R&D Organizations: Some Preliminary Findings by William M. Evan.

Cook, Thomas J. and Frank P. Scioli, Jr., *A Research Strategy for Analyzing the Impacts of Public Policy*. Administrative Science Quarterly, Vol. 17, No. 3 (September 1972): 328-339.

Relates to impact.

An excellent article. Research strategy for measuring policy impacts based on the principles of experimental design is described. The research model, which stresses policy impacts rather than outputs, deals with program objectives, environment articles, events, alternatives, and policy evaluation criteria.

Dean, Burton V. Evaluating, Selecting, and Controlling R&D Projects. American Management Association Research Study 89. New York: American Management Association, 1968.

Relates to input and output.

This research report is based on the results of a questionnaire survey, interviews with corporations, and on a review of relevant literature. Chapters 4 through 7 deal with the following topics: Project Evaluation: Methods and Procedures, Project (selection) Formulas, Project Control, and Projects Completion and Termination.

Dean, Burton V. Operations Research in Research and Development. New York: John Wiley and Sons, Inc., 1963.

Relates to input, process, and output.

Some particularly relevant papers are as follows:

Paper 3, *The Measurement of Value of Scientific Information*, by Miles W. Martin, Jr. The use of citation counts as a method of measuring scientific value are discussed and evaluated. A generalized mathematical model to determine objectively the value and amount of scientific information contained in an article is presented. This is followed by an experimental application of the theory.

Paper 6, *Selection, Evaluation, and Control of Research and Development Projects*, by David B. Hertz and Phillip G. Carlson. On procedures, organization, administration, and criteria to select, evaluate, and control research projects. Although addressed to private business, much of the paper applies to R&D programs in government.

Paper 7, *Studies of Project Selection Behavior in Industry*, by Albert H. Rubenstein. A general discussion of project selection, estimation, and the behavior of individual decision makers. A brief review of related literature is included.

Dean, Burton V. and Meir J. Nishry, *Scoring and Profitability Models for Evaluating and Selecting Engineering Projects*. Operations Research, Vol. 13, No. 4 (July-August 1965): 550-569.

Relates to value analysis.

A paper concerned with the problems of evaluating and selecting engineering projects for new product development. The factors affecting the success of the projects were identified and used in a project scoring model. The scoring model was then used to determine the important factors in the development of a project profitability model. Although primarily directed at industrial research, the method has relevance for non-profit research.

Dean, Burton V., et al. Analysis of the Exploratory Development Project Evaluation Experiment. Case Western Reserve University, Cleveland, Ohio. June 1970. (NTIS publication AD 226 879.)

Relates to value analysis.

This paper presents a description of the methods used and the results obtained in an exploratory project evaluation experiment. The experiment, designed as a modification of the Delphi Method, uses a panel to evaluate research projects. Results of the experiment are discussed in depth. A variety of statistical techniques are used to analyze the data obtained from the project evaluators.

Dorfman, Robert, ed. Measuring Benefits of Government Investments. Papers presented at a conference held November 7-9, 1963. Washington, DC: The Brookings Institution, 1965.

Relates to impact.

The book is composed of seven papers on appraising the benefits that are likely to accrue from proposed public investment projects. *Government Research and Development Programs*, by Frederic M. Scherer is particularly pertinent. Six other papers analyze the benefits of programs in the following areas: outdoor recreation, high school dropouts, civil aviation, urban highways, urban renewal, and syphilis control.

Edwards, Shirley A. and Michael W. McCarrey, *Measuring the Performance of Researchers*. Research Management, Vol. XVI, No. 1 (January 1973): 34-41.

Relates to output.

This review of studies dealing with the measurement of scientific performance reveals that there is very little agreement as to what constitutes scientific output or what measures of output should be used. Methods for determining performance range from rating by peers and supervisors to complex evaluations of the quantity and quality of written output. The paper states that scientific output is multidimensional and that measurement by a number of independent attributes is desirable.

Glass, E.M. Evaluation of R&D Organizations. Washington, DC.: Office of the Director of Defense Research and Engineering. July 31, 1969. (NTIS publication AD 697 343.)

Describes a peer rating experiment to evaluate DOD R&D laboratories. The laboratories are rated by personnel at various organizational levels. The assumption is that the quality of scientific work will be at the same level as the technical reputation of a laboratory in the scientific community. In the future, an attempt will be made to relate performance to organizational characteristics.

Gruber, William H. and Donald G. Marquis. Factors in Technology Transfer. Cambridge, MA.: The M.I.T. Press, 1969.

Relates to output and impact.

A paper, *The Differential Performance of Information Channels in the Transfer of Technology*, by Thomas J. Allen, deals with the importance of various communication channels in the transfer of scientific and engineering research information.

Heyel, Carl, ed. Handbook of Industrial Research Management. 2nd ed. New York: Reinhold Book Company, 1968.

Relates to value analysis.

The most pertinent chapters include the following:

Chapter 7, Establishing Research Projects (dealing with project selection), by W.C. Asbury;
Chapter 12, Executive Direction and Progress Measurement, by William E. Camp; and
Chapter 13, Top Management Reports and Control, by Robert M. Bowie in collaboration with Irwin Goldman, Donato Bracco, and Daniel Lazare.

Joyce, William B., *Organization of Unsuccessful R&D Projects*. IEEE Transactions on Engineering Management, EM-18 (May 1971): 57-65.

Relates to input, process, and output.

This paper discusses the organization of an applied research project, regarding a project as a collection of necessary but potentially successful tasks and subtasks. A procedure for selecting which projects to fund and for estimating expenditures is given. A method for sequencing tasks and minimizing cost is suggested. The method takes into account the possibility that the project may fail before all tasks are completed. The author characterizes the R&D effort as, at least in part, a process of identifying unsuccessful ideas.

Lebanoff, Lazarus, *Total Evaluation of Management Purposes of Engineering and Scientific Tasks*. IEEE Transactions on Engineering Management, EM-13 (June 1966): 110-122.

Relates to process and output.

This excellent article analyzes a system for reporting to management the progress of scientific and engineering projects. The assumption is that the critical elements of information affecting technical projects can be isolated and meaningfully quantified. An analysis procedure was developed to change quantified data into meaningful indicators for management so that research projects could be continuously assessed.

Lipetz, Ben-Ami. The Measurement of Efficiency of Scientific Research. Intermedia, Inc., Carlisle, MA: 1965. Q 180 A1K5, 1965.

Relates to output.

A general, theoretical approach to efficiency and effectiveness measurement in scientific research. The chapters dealing with the inadequacy of current measurement techniques, the validity and limitations of the efficiency concept, and the recurring products of scientific research are most useful. The recurring products of scientific research: description, defining hypotheses, explanation, prediction, and experimental technique, are presented as desirable and measurable output of scientific activity. There is no evidence to suggest that the proposed measurement scheme has been used.

Martino, J.P., *Citation Indexing for Research and Development Management*. IEEE Transactions on Engineering Management, EM-18 (November 1971): 146-151.

Relates to output.

A thorough examination of citation indices as a quantitative measure of the value of a publication. The empirical studies cited show a positive correlation between quality of a publication and its citation rate. Three possible uses of citation indices are given: 1) to evaluate quality of an individual, 2) to evaluate the quality of a group effort, and 3) to determine if researchers are aware of information relevant to their own work. The reader is cautioned that while citation indices offer a tool for R&D managers, other factors are also important.

Molander, Robert C., et al. The Evaluation of Research Supported by a Mission-Oriented Agency Structured Approaches. Volume I: On Measures of Research Effectiveness (Paper P-753). Institute for Defense Analyses, Advanced Research Projects Agency. December 1971. (NTIS publication AD 753 814)

The feasibility of developing structured approaches to prospective and retrospective evaluation of basic research is examined. A search was conducted to an approach which measures scientific disciplines and sub-disciplines according to their relevance to the agency's mission and which measures individual research

projects within the categories against a yardstick of scientific excellence. No successful measures for relevance assessment or for prospective evaluation of individual projects were identified. Retrospective measures for project evaluation were found to be feasible but the problems associated with them are considered too serious to warrant relying on them.

Moore, John R., Jr., and Norman R. Baker, *An Analytical Approach to Scoring Model Design - Application to Research and Development Project Selection.* IEEE Transactions on Engineering Management, EM-16 (August 1969): 90-98.

Relates to value analysis.

This article contains a detailed discussion of an analytical method for scoring model design and for verification. The scoring model compares favorably in statistical accuracy and sensitivity to other more widely accepted models for project evaluation. The following aspects of model development are discussed:

(1) selection of evaluation criteria
(2) development of performance measures
(3) quantification of the research environment
(4) determination of criteria weights
(5) initial model specification
(6) selection of model objectives
(7) initial model verification
(8) complete model specification and verification.

Moravcsik, Michael J., *Measures of Scientific Growth.* Research Policy, 2 (1973), 266-275.

Relates to output and impact.

An excellent discussion of some of the problems of using publications and citations as measures of scientific growth or progress.

National Bureau of Economic Research. The Rate and Direction of Economic Activity: Economic and Social Factors. (A Conference of the Universities --National Bureau Committee for Economic Growth of the Social Science Research Council). Princeton, NJ: Princeton University Press, 1962.

Relates to input, process, output, and impact.

Some of the most pertinent papers include the following:

*Inventive Activity: Problems of Definition and Measurement,* by Simon Kuznets;
*Some Difficulties in Measuring Inventive Activity,* by Barker S. Sanders;

*Predictability of the Costs, Time, and Success of Development,*
by A.W. Marshall and W.H. Herckling; and
*The Decision Making Problem in Development,*
by Burton H. Klein.

Pelz, Donald C. and Frank M. Andrews. Scientists in Organization Productive Climates for Research and Development. New York: John Wiley and Sons, Inc., 1966.

The book examines the relationship between scientific performance and the organization of a laboratory in terms of freedom, communication, diversity, dedication, motivation, satisfaction, similarity, creativity, age, climate, and communication.

Pound, William H., *Research Project Selection: Testing a Model in the Field.* IEEE Transactions on Engineering Management, EM-11 (March 1964): 16-22.

Relates to value analysis.

This article presents the results of a field test of a procedure for evaluating research projects. The procedure is based on an expected value. It considers the following elements: the environment of the problem, the decision maker, the objectives, and the alternatives. The result of the procedure was a ranking of potential projects in terms of their expected values. The resulting ranking of the projects agreed with an intuitive evaluation by the decision makers of the same list of projects, indicating that the model may be useful in the area of research project selection.

Quinn, James Brian. Yardsticks for Industrial Growth. New York: The Ronald Press, 1959.

Relates to input, output, and value analysis.

Discusses the evaluation of scientific research efforts in profit motivated enterprises in the United States. Some of the more general chapters relevant to non-profit motivated research are:

Chapter 2, *A Survey of Current Evaluation Techniques;*
Chapter 3, *Structuring the Problem;*
Chapter 4, *The Technical Evaluation of Research;*
Chapter 8, *Evaluating the Management of the Research and Development Program,* and
Chapter 9, *Summary Segmental Evaluations.*

Chapters 5 through 7 deal, respectively, with the economic evaluation of offensive research, defense research and fundamental research, and discusses specific techniques and formulas for industrial research.

Quinn, James L., et al. A Categorization of the Methods and Techniques of Measuring Industrial R&D Activities. Air University, Wright-Patterson Air Force Base, OH. February 1971. (NTIS publication AD 727 023.)

Discusses a variety of quantitative and qualitative analyses employed in measuring and evaluating R&D activities and classifies the methods in terms of pre-appraisal, in process measurement, and post evaluation. The study concludes that quantitative methods should be used whenever appropriate.

Roberts, Edward B. The Dynamics of Research and Development. New York: Harper and Row, Publishers, 1964.

Relates to input, process, and output.

In the first part of this excellent book, Roberts presents a general theory of research and development with chapters addressed to the following topics: structure of projects, perception of product value, project effort and cost, funding, manpower, control, and determinants of project life cycle. Part II discusses the results of a large number of studies about the general theory of R&D project dynamics that were presented in Part I. Numerous quantitative techniques and models are discussed. The chapters most pertinent to productivity measurement appear to be:

Chapter 2, *Perception of Product Value;*
Chapter 3, *Evaluation of Project Cost and Effort;*
Chapter 6, *Control of Research and Development Progress; and*
Chapter 7, *Key Determinants of Project Life Cycle.*

Roman, Daniel. D., *Project Management Recognizes R&D Performance.* Academy of Management Journal, Vol. 7, No. 1 (March 1964), 7-20.

Relates to input, process, and output.

Discusses the need for R&D project evaluations and proposes a model to serve as a master project plan. The project plan is designed to include the following elements: work description, budget allocation, project revision, status reporting, project support assignments and a work log.

Roman, Daniel D. Research and Development Management: The Economics and Administration of Technology. New York: Appleton-Century-Crofts, 1968.

Relates to input, process, and output.

Chapters 10, 11, 13, 14, 15, and 16, dealing with project selection, project management, R&D estimating, costing and budgeting, planning, control, and evaluation and measurement, are particularly useful.

Chapter 16, Evaluation and Measurement, discusses performance standards, appraisal methods, and performance criteria. The discussion covers the public and private sectors and includes qualitative as well as quantitative factors.

Rubenstein, Albert H., *Economic Evaluation of Research and Development: A Brief Survey of Theory and Practice*. The Journal of Industrial Engineering, Vol. 17, No. 11 (November 1966): 615-620.

The article presents a brief overview of the state of the art in economic evaluation of R&D. It describes current research at Northwestern University aimed at designing a "real-time" computer aided information system for project selection, review, and evaluation.

Seiler, Robert E. Improving the Effectiveness of Research and Developm New York: McGraw-Hill Book Company, 1965.

Relates to input, process, and output.

Although addressed to corporate industrial research and development, much of the book is relevant to government research. Some of the chapters that are particularly useful include:

Chapter 4, *The Formulation and Application of Research Objectives;*
Chapter 10, *Selection of Projects - Key to Research Effectiveness;*
Chapter 11, *Selection of Projects - The Research Proposal;*
Chapter 12, *Selection of Projects - Quantitative Methods; and*
Chapter 14, *Techniques for an Overall Evaluation of the Research Effort.*

Seyfried, W.D., *The Evaluation of Research*. American Management Association Report No. 76, 1963, 215-221.

The author provides a good conceptual overview of research evaluation and its relationship to planning and organization. The major themes addressed briefly in the article are the purpose of evaluation, the responsibility for evaluation, and methods for evaluation.

Souder, William E., *Experiences with an R&D Project Control Model*. IEEE Transactions on Engineering Management, EM-15 (March 1968): 39-49.

R&D project cost control should relate expenditures to achievement as well as to time elapsed. A theoretical model relating to cost, achievement, and time is described. The model was tested on two pilot experiments involving chemical R&D projects. The model provided early warning of project failures and pinpointed the forces affecting the failures. It gave a detailed analysis of the achievement per dollar spent. The author concludes that, generally, if an R&D project shows a high frequency of large revisions in achievement assessment over time producing long plateaus of low achievement, there is a high probability of project failure.

Trozzo, C.L. Description and Critique of Quantitative Methods for the Allocation of Exploratory Development Resources (Paper P-731). Institute for Defense Analysis, Science and Technology Division. May 1972. (NTIS publication AD 753 817.)

The paper analyzes ten methods that have been used as proposals for planning the allocation of resources among projects within the Exploratory Development category of the Defense Research, Development, Test, and Evaluation Program. Each method is described within a general framework of planning methods. A comparative analysis is made of the strengths and weaknesses of these methods.

U.S. National Science Foundation. Technology in Retrospect and Critical Events in Science. Two volumes. Prepared for the NSF by the Illinois Institute of Technology Research Institute, December 15, 1968

Relates to input, process, output, and impact.

A retrospective tracing of key events which led t[illegible] of major technological innovations. Volume 1 cont[illegible] overall analyses and conclusions and case studies of following innovations: magnetic ferrites, video tape recorders, oral contraceptives, electron microscopes, and matrix isolation. Volume 2 gives more detailed technical information about the various events relating to these innovations.

U.S. National Science Foundation. Interactions of Science and Technology in the Innovative Process: Some Case Studies. Prepared for NSF. Columbus, OH: Battelle, March 19, 1973.

Relates to input, process, output, and impact.

This report is an attempt to understand the innovative process by retrospective studies documenting historically the significant events in ten technological innovations of high social impact. Many factors and generalizations relating to the innovative process are discussed. Detailed base studies of innovations include the heart pacemaker, hybrid grains, electrophotography, and economic input-output analysis.

Vollmer, H.R., *Evaluating Two Aspects of Quality in Research Program Effectiveness*. Research Program Effectiveness. Edited by M.C. Yovits, et al. New York: Gordon and Breach, Science Publishers, Inc., 1966. 147-167.

Relates to output and impact.

Studies of research organization and effectiveness by the Office of Aerospace Research (U.S. Air Force) and the Office of Civil Defense suggest the feasibility of two measures of the quality of research program outputs. The first measure focuses on the source of publication of research fundings, while the second deals with number of citations appearing in subsequent literature, i.e., the amount of recognition achieved. Interfacility comparisons of sources of publications and higher degree of consistency between the two comparisons suggest that the two measures are feasible indicators of the quality of basic scientific research output of a facility or a program.

Walters, J.E. Research Management: Principles and Practice. Washington, DC: Spartan Books, 1965.

Relates to input, process, and output.

Although the book covers the status of research and development in the United States and is not addressed specifically to Federal R&D programs, many of the discussions are applicable to any R&D effort. Chapters 5, 9, 11, and 12, dealing with project formulation, performance, coordination and evaluation, and audit, are pertinent to productivity measurement.

Weiss, Carol H. Evaluation Research Methods for Assessing Program Effectiveness. Englewood Cliff, NJ: Prentice-Hall, Inc., 1972.

Relates to output, impact, and value analysis.

This book deals with the application of research methods to the evaluation of social programs. Written for use on a basic text on evaluation research, the central theme of the work concerns the application of social research methods and tools within the context of social program evaluation. A very good bibliography is included.

Williams, D.J., *A Study of a Decision Model for R&D Project Selection*. Operational Research Quarterly, Vol. 20, No. 3 (1969): 361-373.

Relates to value analysis.

This paper reports on a study made during 1966 at the Bristol Works of the British Aircraft Corporation on the evaluation and selection of R&D projects. After identifying the objectives and criteria used in the selection process, an attempt was made to establish the relative importance and interaction among the various factors. The intent of the study was to derive a model based on a project scoring system using weighted sums of factor scores. Although data were insufficient to develop such a model, it was shown that to represent reality the interdependence of the selection criteria and objectives must be explicitly established.

NBS-114A (REV. 7-73)

| U.S. DEPT. OF COMM. BIBLIOGRAPHIC DATA SHEET | 1. PUBLICATION OR REPORT NO. NBS TN-890 | 2. Gov't Accession No. AD-A106 487 | 3. Recipient's Accession No. |
|---|---|---|---|
| 4. TITLE AND SUBTITLE<br>PRODUCTIVITY MEASUREMENT IN R&D: Productivity Measurement Experiment (PROMEX) in Selected Research and Development Programs at the National Bureau of Standards | | | 5. Publication Date<br>December 1975 |
| | | | 6. Performing Organization Code |
| 7. AUTHOR(S) John T. Hall and Roger A. Dixon | | | 8. Performing Organ. Report No. |
| 9. PERFORMING ORGANIZATION NAME AND ADDRESS<br>NATIONAL BUREAU OF STANDARDS<br>DEPARTMENT OF COMMERCE<br>WASHINGTON, D.C. 20234 | | | 10. Project/Task/Work Unit No. |
| | | | 11. Contract/Grant No. |
| 12. Sponsoring Organization Name and Complete Address *(Street, City, State, ZIP)*<br>Same as 9. | | | 13. Type of Report & Period Covered<br>Final |
| | | | 14. Sponsoring Agency Code |

15. SUPPLEMENTARY NOTES

16. ABSTRACT *(A 200-word or less factual summary of most significant information. If document includes a significant bibliography or literature survey, mention it here.)*


*This report describes an experiment in productivity measurement conducted at the National Bureau of Standards. The experiment concludes that no matter how sophisticated the analysis and synthesis processes become, statistical counts of* output *media (e.g., publications, citations, invited talks) will not serve as reliable measures of R&D productivity.*

*The conduct of the experiment included a work sampling study, a communications study, an output analysis, a value analysis approach to developing criteria for selection and evaluation of programs, construct of a rating system for evaluation of programs, and construction of a model of the R&D process.*


17. KEY WORDS *(six to twelve entries; alphabetical order; capitalize only the first letter of the first key word unless a proper name; separated by semicolons)*

Impact; objectives, output; performance measurement; production indices; product[illegible] measurement.

| 18. AVAILABILITY [X] Unlimited | 19. SECURITY CLASS (THIS REPORT) UNCLASSIFIED | 21. NO. OF PAGES 52 |
|---|---|---|
| [ ] For Official Distribution. Do Not Release to NTIS | | |
| [X] Order From Sup. of Doc., U.S. Government Printing Office Washington, D.C. 20402, SD Cat. No. C13. 46:890 | 20. SECURITY CLASS (THIS PAGE) UNCLASSIFIED | 22. Price $1.10 |
| [ ] Order From National Technical Information Service (NTIS) Springfield, Virginia 22151 | | |

USCOMM-DC 29042-P74

# NBS TECHNICAL PUBLICATIONS

## PERIODICALS

JOURNAL OF RESEARCH reports National Bureau of Standards research and development in physics, mathematics, and chemistry. It is published in two sections, available separately:

• **Physics and Chemistry (Section A)**

Papers of interest primarily to scientists working in these fields. This section covers a broad range of physical and chemical research, with major emphasis on standards of physical measurement, fundamental constants, and properties of matter. Issued six times a year. Annual subscription: Domestic, $17.00; Foreign, $21.25.

• **Mathematical Sciences (Section B)**

Studies and compilations designed mainly for the mathematician and theoretical physicist. Topics in mathematical statistics, theory of experiment design, numerical analysis, theoretical physics and chemistry, logical design and programming of computers and computer systems. Short numerical tables. Issued quarterly. Annual subscription: Domestic, $9.00; Foreign, $11.25.

DIMENSIONS/NBS (formerly Technical News Bulletin)—This monthly magazine is published to inform scientists, engineers, businessmen, industry, teachers, students, and consumers of the latest advances in science and technology, with primary emphasis on the work at NBS. The magazine highlights and reviews such issues as energy research, fire protection, building technology, metric conversion, pollution abatement, health and safety, and consumer product performance. In addition, it reports the results of Bureau programs in measurement standards and techniques, properties of matter and materials, engineering standards and services, instrumentation, and automatic data processing.

Annual subscription: Domestic, $9.45; Foreign, $11.85.

## NONPERIODICALS

**Monographs**—Major contributions to the technical literature on various subjects related to the Bureau's scientific and technical activities.

**Handbooks**—Recommended codes of engineering and industrial practice (including safety codes) developed in cooperation with interested industries, professional organizations, and regulatory bodies.

**Special Publications**—Include proceedings of conferences sponsored by NBS, NBS annual reports, and other special publications appropriate to this grouping such as wall charts, pocket cards, and bibliographies.

**Applied Mathematics Series**—Mathematical tables, manuals, and studies of special interest to physicists, engineers, chemists, biologists, mathematicians, computer programmers, and others engaged in scientific and technical work.

**National Standard Reference Data Series**—Provides quantitative data on the physical and chemical properties of materials, compiled from the world's literature and critically evaluated. Developed under a world-wide program coordinated by NBS. Program under authority of National Standard Data Act (Public Law 90-396).

NOTE: At present the principal publication outlet for these data is the Journal of Physical and Chemical Reference Data (JPCRD) published quarterly for NBS by the American Chemical Society (ACS) and the American Institute of Physics (AIP). Subscriptions, reprints, and supplements available from ACS, 1155 Sixteenth St. N. W., Wash. D. C. 20056.

**Building Science Series**—Disseminates technical information developed at the Bureau on building materials, components, systems, and whole structures. The series presents research results, test methods, and performance criteria related to the structural and environmental functions and the durability and safety characteristics of building elements and systems.

**Technical Notes**—Studies or reports which are complete in themselves but restrictive in their treatment of a subject. Analogous to monographs but not so comprehensive in scope or definitive in treatment of the subject area. Often serve as a vehicle for final reports of work performed at NBS under the sponsorship of other government agencies.

**Voluntary Product Standards**—Developed under procedures published by the Department of Commerce in Part 10, Title 15, of the Code of Federal Regulations. The purpose of the standards is to establish nationally recognized requirements for products, and to provide all concerned interests with a basis for common understanding of the characteristics of the products. NBS administers this program as a supplement to the activities of the private sector standardizing organizations.

**Federal Information Processing Standards Publications (FIPS PUBS)**—Publications in this series collectively constitute the Federal Information Processing Standards Register. Register serves as the official source of information in the Federal Government regarding standards issued by NBS pursuant to the Federal Property and Administrative Services Act of 1949 as amended, Public Law 89-306 (79 Stat. 1127), and as implemented by Executive Order 11717 (38 FR 12315, dated May 11, 1973) and Part 6 of Title 15 CFR (Code of Federal Regulations).

**Consumer Information Series**—Practical information, based on NBS research and experience, covering areas of interest to the consumer. Easily understandable language and illustrations provide useful background knowledge for shopping in today's technological marketplace.

**NBS Interagency Reports (NBSIR)**—A special series of interim or final reports on work performed by NBS for outside sponsors (both government and non-government). In general, initial distribution is handled by the sponsor; public distribution is by the National Technical Information Service (Springfield, Va. 22161) in paper copy or microfiche form.

---

Order NBS publications (except NBSIR's and Bibliographic Subscription Services) from: Superintendent of Documents, Government Printing Office, Washington, D.C. 20402.

## BIBLIOGRAPHIC SUBSCRIPTION SERVICES

**The following current-awareness and literature-survey bibliographies are issued periodically by the Bureau:**
Cryogenic Data Center Current Awareness Service

A literature survey issued biweekly. Annual subscription: Domestic, $20.00; foreign, $25.00.

**Liquefied Natural Gas.** A literature survey issued quarterly. Annual subscription: $20.00.

**Superconducting Devices and Materials.** A literature survey issued quarterly. Annual subscription: $20.00. Send subscription orders and remittances for the preceding bibliographic services to National Technical Information Service, Springfield, Va. 22161.

**Electromagnetic Metrology Current Awareness Service** Issued monthly. Annual subscription: $24.00. Send subscription order and remittance to Electromagnetics Division, National Bureau of Standards, Boulder, Colo. 80302.