DEPARTMENT
OF
COMMERCE

MISCELLANEOUS
PUBLICATIONS
OF THE
NATIONAL
BUREAU
OF
STANDARDS

NOS. 247-255

Proceedings of the 1962 Standards Laboratory Conference

United States Department of Commerce
National Bureau of Standards
Miscellaneous Publication 248

THE NATIONAL BUREAU OF STANDARDS

Functions and Activities

The functions of the National Bureau of Standards are set forth in the Act of Congress, March 3, 1901, as amended by Congress in Public Law 619, 1950. These include the development and maintenance of the national standards of measurement and the provision of means and methods for making measurements consistent with these standards; the determination of physical constants and properties of materials; the development of methods and instruments for testing materials, devices, and structures; advisory services to government agencies on scientific and technical problems; invention and development of devices to serve special needs of the Government; and the development of standard practices, codes, and specifications. The work includes basic and applied research, development, engineering, instrumentation, testing, evaluation, calibration services, and various consultation and information services. Research projects are also performed for other government agencies when the work relates to and supplements the basic program of the Bureau or when the Bureau's unique competence is required. The scope of activities is suggested by the listing of divisions and sections on the inside of the back cover.

Publications

The results of the Bureau's research are published either in the Bureau's own series of publications or in the journals of professional and scientific societies. The Bureau publishes three periodicals available from the Government Printing Office: The Journal of Research, published in four separate sections, presents complete scientific and technical papers; the Technical News Bulletin presents summary and preliminary reports on work in progress; and the Central Radio Propagation Laboratory Ionospheric Predictions provides data for determining the best frequencies to use for radio communications throughout the world. There are also five series of nonperiodical publications: Monographs, Applied Mathematics Series, Handbooks, Miscellaneous Publications, and Technical Notes.

A complete listing of the Bureau's publications can be found in National Bureau of Standards Circular 460, Publications of the National Bureau of Standards, 1901 to June 1947 ($1.25), and the Supplement to National Bureau of Standards Circular 460, July 1947 to June 1957 ($1.50), and Miscellaneous Publication 240, July 1957 to June 1960 (includes Titles of Papers Published in Outside Journals 1950 to 1959) ($2.25); available from the Superintendent of Documents, Government Printing Office, Washington 25, D.C.

UNITED STATES DEPARTMENT OF COMMERCE • Luther H. Hodges, *Secretary*
NATIONAL BUREAU OF STANDARDS • A. V. Astin, *Director*

Proceedings of the 1962 Standards Laboratory Conference

Presented by the National Conference of Standards Laboratories, August 8, 9, 10, 1962, National Bureau of Standards, Boulder Laboratories, Boulder, Colorado

National Bureau of Standards Miscellaneous Publication 248

Issued August 16, 1963

For sale by the Superintendent of Documents, U.S. Government Printing Office, Washington, D.C., 20402
Price $1.75

1962 STANDARDS CONFERENCE COMMITTEE

Conference Chairman

Dr. Allen V. Astin, Director National Bureau of Standards

Administrative

Chairman: Harvey W. Lance, NBS
Secretary: Alfred E. Hess, NBS
Publicity: Jack R. Craddock, NBS
Finance: Herbert D. Stansell, NBS
Publication: Wilbert F. Snyder, NBS
Local Arrangements: James F. Brockman, NBS

Technical Program

Chairman: Charles E. Johnson, The Boeing Company
Members: William G. Amey, Leeds and Northrup Company
Leon W. Bowen, The Boeing Company
Jerry L. Hayes, Bureau of Naval Weapons
K. G. Overbury, Physical & Electric Standards
S. C. Richardson, General Electric Company
L. R. Wallace, International Business Machines
William A. Wildhack, NBS
Lloyd B. Wilson, Sperry Gyroscope Company
A. J. Woodington, General Dynamics/Astronautics

GENERAL COMMITTEE OF THE NATIONAL CONFERENCE OF STANDARDS LABORATORIES

1962

Lloyd B. Wilson, Sperry Gyroscope Company, Chairman
Charles E. Johnson, The Boeing Company, Vice Chairman
Harvey W. Lance, National Bureau of Standards, Boulder, Corresponding Secretary
Charles White, Avco, R.A.D., Recording Secretary and Treasurer
William G. Amey, Leeds and Northrup Company
Herbert Barnhart, General Electric Company
H. C. Biggs, Sandia Corporation
Melvin Fruechtenicht, U. S. Army Ordnance
Jerry L. Hayes, Bureau of Naval Weapons
Wallace L. Horton, U. S. Air Force
Peter Joeschke, Autonetics
William A. Wildhack, National Bureau of Standards, Washington

The results of the elections held during the 1962 Conference are as follows:

Charles E. Johnson, The Boeing Company, Seattle, Chairman
Harvey W. Lance, National Bureau of Standards, Boulder, Vice Chairman
Ivan G. Easton, General Radio, Corresponding Secretary
William G. Amey, Leeds and Northrup Company
Herbert D. Barnhart, General Electric Company
H. C. Biggs, Sandia Corporation, Alburquerque
Jerry L. Hayes, Bureau of Naval Weapons, Pomona
Peter A. Joeschke, NAA-Autonetics, Downey
Charles E. White, Avco R.A.D., Recording Secretary and Treasurer (second year of two year term.)

Library of Congress Catalog Card Number: 63-60068

FOREWORD

More than 600 persons from approximately 200 industrial laboratories, and other organizations concerned with measurement standards, convened at the Boulder Laboratories of NBS on August 8-10, 1962, for the first national meeting of the National Conference of Standards Laboratories.

The National Conference of Standards Laboratories provides a means by which the country's standards laboratories may cooperate in generating and disseminating useful information relating to calibration techniques and to the operation of standards laboratories. NBS has therefore encouraged the organization and activities of the Conference, and will continue to provide assistance in mutually useful activities, as valuable supplements to the Bureau's work in disseminating accuracy of measurement throughout science and industry. The publication of the Proceedings of this Conference, containing papers presented at the national meeting, is one example of the Bureau's cooperation.

Most of the papers presented at the meeting are published in this volume. Primary responsibility for their technical content must rest, of course, with the individual authors and their organizations.

A. V. Astin, Director

WELCOMING REMARKS

Dr. A. V. Astin, Director of the National Bureau of Standards and General Chairman of the Standards Laboratory Conference of NCSL, and Dr. F. W. Brown, Director, NBS Boulder Laboratories, welcomed the attendees.

Dr. Astin called attention to the significance of standards laboratories with respect to the country's technological economy. He expressed his pleasure at the large attendance and indicated the great interest of the National Bureau of Standards in the work of the Conference as an organization devoted to the interests of standards laboratories. He called attention to two important roles for NCSL: first, acquisition or exchange of information of interest to participating laboratories in many areas in which NBS is not involved so directly; and second, cooperation with NBS in more effectively disseminating information on accurate measurement.

Contents

Session 1. National Bureau of Standards Service to Industry

INTRODUCTORY REMARKS BY SESSION CHAIRMAN

W. A. Wildhack

It is indeed an honor to have this opportunity to open the technical sessions of the 1962 National Conference of Standards Laboratories with a report of the National Bureau of Standards activities in the area of particular interest to other calibration laboratories. In his welcoming remarks Dr. Astin expressed the pleasure of the NBS staff in being hosts to the Conference here at the Boulder Laboratories.

I would add that we in NBS share with you, as NCSL participants, some considerable gratification that more than 550 persons, representatives from at least 200 laboratories, are here for the Conference. This wide participation not only confirms the practical value of NCSL, but augurs well for its continuing success. We in NBS would like to extend our congratulations to the NCSL officers and committees for their achievement in organizing this Conference.

For the larger number of present attendees, it may be recalled that the NCSL arose from a suggestion, or a hope, expressed by Mr. Harvey Lance during his talk at the Conference on Precision Electromagnetic Measurements held in June 1960 here at NBS Boulder. More than 150 of the 800 persons attending that conference were sufficiently interested to assemble an hour earlier the following day to discuss this suggestion for an organization of standards laboratories. This discussion led to a request for the General Committee of the CPEM to name a special ad hoc committee to explore further the needs and to make recommendations for the appropriate role and structure of an organization devoted to the interests of standards laboratories. With the cooperation of the Instrument Society of America, the ad hoc committee, under the chairmanship of H. C. Biggs, organized a session at the 1961 ISA Conference in Los Angeles; at that time there was established a continuing General Committee to arrange a series of national conferences and to carry on projects of value and interest to standards laboratories through committee action. This Conference we now begin is evidence that the committees have been diligently pursuing that goal, under the active leadership of Mr. Lloyd Wilson.

The program for this morning is devoted to a presentation of current and planned activities of the National Bureau of Standards in attempting to provide adequate measurement services to other standards laboratories.

Paper 1.1. The Measurement Services Program of the National Bureau of Standards

W. A. Wildhack*

This paper provides an outline of the content of the National Bureau of Standards measurement services program, current tasks, work underway and planned, and relationships with other organizations and industrial and governmental programs.

1. Introduction

Because most people in science and industry, even those who deal with NBS, are not aware of many of the Bureau's functions, it may be desirable to give a brief outline of some of the other activities of NBS before considering the measurement services in more detail.

On its inception in 1901, NBS had the statutory functions of: custody of standards, and the comparison of these standards with those of science and industry; the testing and calibration of standard measuring apparatus; the solution of problems which arise in connection with standards; the determination of physical constants and properties of materials; and the dissemination of information concerning standards, methods of measurement, and properties of materials.

measurement, and the provision for means of making measurements consistent with those standards; determination of physical constants and properties of materials; development of methods for testing materials, mechanisms, and structures, and the making of such tests as may be necessary, particularly for Government agencies; cooperation in the establishment of standard practices, incorporated in codes and specifications; advisory service to Government agencies on scientific and technical problems; and invention and development of devices to serve special needs of the Government.

A wide diversity of work in these broad categories is carried on by 23 technical Divisions. The present technical Division and Section structure of NBS[1] is as follows:

NBS WASHINGTON

SCIENTIFIC DIVISIONS AND SECTIONS

Electricity

Resistance and Reactance
Electrochemistry
Electrical Instruments
Magnetic Measurements
Dielectrics
High Voltage

Heat

Temperature Physics
Heat Measurements
Cryogenic Physics
Equation of State
Statistical Physics

Analytical and Inorganic Chemistry

Pure Substances
Spectrochemistry
Solution Chemistry
Standard Reference Materials
Applied Analytical Research
Crystal Chemistry

Polymers

Macromolecules: Synthesis and Structure
Polymer Chemistry
Polymer Physics
Polymer Characterization
Polymer Evaluation and Testing
Applied Polymer Standards and Research
Dental Research

Metrology

Photometry and Colorimetry
Refractometry
Photographic Research
Length
Engineering Metrology
Mass and Scale
Volumetry and Densimetry

Radiation Physics

X-ray
Radioactivity
Radiation Theory
High Energy Radiation
Radiological Equipment
Nucleonic Instrumentation
Neutron Physics

Mechanics

Sound
Pressure and Vacuum
Fluid Mechanics
Engineering Mechanics
Rheology
Combustion Controls

Metallurgy

Engineering Metallurgy
Microscopy and Diffraction
Metal Reactions
Metal Physics
Electrolysis and Metal Deposition

---

*Associate Director, National Bureau of Standards, Washington, D.C.

[1] As of September 1, 1962.

Inorganic Solids

Engineering Ceramics
Glass
Solid State Chemistry
Crystal Growth
Physical Properties
Crystallography

Applied Mathematics

Numerical Analysis
Computation
Statistical Engineering
Mathematical Physics
Operations Research

Atomic Physics

Spectroscopy
Infrared Spectroscopy
Far Ultraviolet Physics
Solid State Physics
Electron Physics
Plasma Spectroscopy

Physical Chemistry

Thermochemistry
Surface Chemistry
Organic Chemistry
Molecular Spectroscopy
Elementary Processes
Mass Spectrometry
Photochemistry and Radiation Chemistry

Building Research

Structural Engineering
Fire Research
Mechanical Systems
Organic Building Materials
Codes and Safety Standards
Heat Transfer
Inorganic Building Materials
Metallic Building Materials

Data Processing Systems

Components and Techniques
Computer Technology
Measurements Automation
Engineering Applications
Systems Analysis

Instrumentation

Engineering Electronics
Electron Devices
Electronic Instrumentation
Mechanical Instruments
Basic Instrumentation

Office of Weights and Measures

- - -

BOULDER DIVISIONS

Cryogenic Engineering

Cryogenic Equipment
Cryogenic Processes
Properties of Materials
Cryogenic Technical Services

Ionosphere Research and Propagation

Low Frequency and Very Low Frequency Research
Ionosphere Research
Prediction Services
Sun-Earth Relationships
Field Engineering
Radio Warning Services
Vertical Soundings Research

Radio Propagation Engineering

Data Reduction Instrumentation
Radio Noise
Tropospheric Measurements
Tropospheric Analysis
Propagation-Terrain Effects
Radio Meteorology
Lower Atomosphere Physics

Upper Atmosphere and Space Physics

Upper Atmosphere and Plasma Physics
High Latitude Ionospheric Physics
Ionosphere and Exosphere Scatter
Airglow and Aurora
Ionospheric Radio Astronomy

Radio Physics

Radio Broadcast Service
Radio and Microwave Materials
Atomic Frequency and Time Interval Standards
Radio Plasma
Millimeter-Wave Research

Radio Systems

- Applied Electromagnetic Theory
- High Frequency and Very High Frequency Research
- Frequency Utilization
- Modulation Research
- Antenna Research
- Radiodetermination

Circuit Standards

- Coordinator Calibration Service
- High Frequency Electrical Standards
- High Frequency Calibration Services
- High Frequency Impedance Standards
- Microwave Calibration Services
- Microwave Circuit Standards
- Low Frequency Calibration Services

The total NBS staff is about 3,500; 2,400 in Washington and 1,100 in Boulder.

Many divisions and sections carry on a variety of projects in research and development on NBS funds, or in cooperation with other agencies, in one or more of the six areas, and any one section may or may not perform any calibrations for the public or for other government agencies. It will be noted that a number of the NBS divisions are doing research and development generally directed to some area of technology or applied science in which measurement standards are not the primary object--even though improvement in techniques or measurement generally is one of the objectives.

2. Budget and Organization

In the Bureau's budget for fiscal year 1962, about $24,000,000 came from the direct appropriation for research and technical services, $29,000,000 from funds appropriated for construction and facilities (mainly for new laboratories at Gaithersburg, Md.), and $15,000,000 from payments by other agencies for specific research and development tasks. In addition, $2,100,000 was received in fees charged for calibration services and $500,000 from the sale of samples of standard reference materials.

While the income for calibration and reference materials is small as compared to the total budget, it must be noted that research on new standards, on how to calibrate, on how to measure, is mostly paid for from appropriated funds. The fees collected from the public are set to cover *only* the approximate cost of the repetitive work of the program, *not* the research on which it is based.

It should be clear from the names of the Bureau's divisions and sections that NBS is looked to for many things by many people and that our efforts must be spread to cover the interests of many groups besides the standards laboratories. Nevertheless, as most of you know, we have been working for several years now to expand and strengthen our calibration services. In fiscal year 1962 this program received increased impetus from a supplemental appropriation of $1.5 million which Congress provided for use in attacking some of the measurement problems pointed out in a survey by the Aerospace Industries Association. It is also gratifying to note that $690,000 has been provided in fiscal year 1963 for the design of a new building for the NBS Radio Standards Laboratory here at Boulder.

3. NBS Measurement Services

3.1. Calibration

One of the major services that NBS provides to the technical public is the calibration of laboratory standards[2], a service that most participants in NCSL are familiar with. At present, we provide well over 100 different calibrations, ranging from Abbe refractometers to X-ray measuring devices. Efforts are constantly being made to improve and extend these services, and to eliminate those which are no longer needed because of advances in the state of the art. For example, an ultrasonic thermometer is being used to establish a temperature scale in the 4 to 14 °K range, and will ultimately provide the basis for the calibration of germanium resistance thermometers. On the other hand, the Bureau discontinued the certification of unsaturated standard cells in 1961, when it was apparent that the demand for such services could be met by other calibration groups.

3.2. Reference Materials

Another part of the Bureau's measurement services program is the provision of standard materials[3] --materials whose physical or chemical properties are highly characterized and which can be used by other laboratories for calibrating instruments or determining the properties of products by comparison. Over 500 such standards are available at present, including steels, irons, brasses, and other metals for spectroscopic standards, radionuclides of known emission rates, pure metals of known freezing point for use in thermometry, *p*H standards, glass spheres for calibration of sieves, and many, many others. In response to the demands of our growing technology, new reference materials are constantly being provided, such as the recently issued samples for glass viscosity standards, and materials which can be adequately handled elsewhere are discontinued.

---

[2] Test Fee Schedules, which list all of the NBS calibration services, may be obtained by writing to the Test Administration Section, National Bureau of Standards, Washington 25, D.C.

[3] A complete listing of all standard samples is contained in NBS Misc. Publ. 241, Standard Materials, available from the Superintendent of Documents, U.S. Government Printing Office, Washington 25, D.C., 30¢.

3.3. Testing

The Bureau's work in testing is mainly for other Government agencies, and testing services are available to the general public only under special circumstances. Quantity testing is restricted to a few items. If a specific problem arises where a particular test is required, or for that matter a particular calibration service, NBS can often provide helpful advice and cooperation.

3.4. Other Bureau Services

The Bureau also provides a wide variety of other services to the technical community. Of these, perhaps the best known, and the most widely available, are the standard frequency and time broadcasts of WWV, WWHV, and more recently, WWVL and WWBV. These around-the-clock broadcasts provide the user with standard radio and audiofrequencies, standard musical pitch, standard time intervals, time signals, and radio propagation forecasts. The precision of these broadcasts has benefited from the development of atomic frequency standards, and the NBS stations are controlled by, or referred to, the cesium beam atomic clocks maintained at Boulder.

3.5. Publications

A variety of technical publications are issued by the Bureau to make the results of its research, and information on measurement techniques, widely available. Such publications as Precision Measurement and Calibration (NBS Handbook 77), a 3-volume compilation of Bureau papers in the field of measurement; the Journal of Research, especially Section C devoted to Engineering and Instrumentation; and the Technical News Bulletin, which contains feature articles devoted to standards and calibrations, serve to provide access to the Bureau's work in the measurement field. The Bureau also publishes data on the properties of materials, or of matter, which would be useful for calibration, or interpretation of results, by others. A recent example is the publication of the absolute transition probabilities of 25,000 spectral lines, information which will be useful to astrophysicists and plasma arc experimenters.

3.6. Visits

The Bureau also maintains an "open-door" policy, providing an opportunity for senior people representing instrument manufacturers or users to come to NBS, talk with our measurement people, and see what we do and why. Such visitors, and we have many of them, not only come away with new information but often make NBS aware of specific needs in their particular field.

3.7. Cooperative Activities

NBS sponsors, or cosponsors with other technical groups, a variety of technical meetings which are of general interest to measurement personnel. Some examples are the 1961 Symposium on Temperature, the annual Conference on Weights and Measures, and the biennial Conferences on Precision Electromagnetic Measurements. The Bureau also cooperates extensively with technical societies in the formulation of standard test methods and specifications.

3.8. Statistical Analyses

The interpretation of a series of observations recorded during the calibration procedure calls for the proper application of appropriate statistical techniques. As estimates of the uncertainty of a calibration take a wide and often bewildering form, depending on the source, the Bureau is seeking to establish a better understanding and general agreement on procedures which will provide a rational basis for the expression of such uncertainties. The use of a common approach to the problem will help to provide, for all echelons of calibration or measurement, more meaningful answers to the oft-asked question--"plus or minus what?" Two Bureau statisticians, Drs. Eisenhart and Youden, are presenting papers at this meeting on aspects of this matter.

3.9. Traceability

A problem which concerns many NCSL members, which we are keenly aware of but not directly involved with, is that of the interpretation, for contractual purposes, of that uncertain and "untraceable" term: "traceability." This is another concept which means many things to many people. The Bureau, of course, has no direct responsibility for defining what is meant by this term as used by others. This rests with those who require "traceability" as part of their contractual arrangements.

In a discussion at the NCSL Workshop on Measurement Agreement last January, it was brought out that, without further definition, the meaning of the term traceability is necessarily indefinite as applied to relationships between calibrations by NBS and measurement activities of manufacturers and suppliers. Traceability is not given any special meaning by the National Bureau of Standards, and information as to possible special meanings of the term in military procurement activities cannot, of course, be supplied by NBS. Where questions arise, they should be directed to the contracting agency.

The difficulty is that practically any measuring instrument has been, at some time, calibrated or checked against some other instrument device which has been, at some time, calibrated or checked against some one of a series terminating with an NBS standard. Thus, all measuring devices are, in this sense, "traceable," but the term *as thus used,* carries no connotation of the accuracy, nor the date, of any one of the calibrations or checks in the series.

In consultation with military procurement officials, we have proposed the substitution of more exact and meaningful terms; the term "traceability" will probably be with us for quite a while--but in combination, we hope, with other terms which add more meaning.

4. Measurement Seminars

The increased demand for accuracy has brought attention to the need for a more direct flow of information from NBS to other measurement personnel. Accordingly, the Bureau is giving attention to providing more information on calibration techniques, not only by the issuance of more publications bearing on this area but also by arrangement of a series of measurement seminars. These seminars, to be held both here at Boulder and at Washington, would provide senior technical people from industry and other government agencies with the opportunity to see what we do, how we do it, and why. As visualized, these seminars would extend from a few days to a week or two, and would be held not only in conference rooms but also, where practical, in the laboratory.

The NCSL can help to put these seminars on a priority basis by letting us know in which areas there exists the most immediate need for information on precision techniques. With this information, we can better set up a schedule for such seminars.

5. Planning NBS Program To Provide Adequate Standards

It would greatly simplify our problems if nobody really needed to make any measurement with more accuracy than is consistent with the calibrations we can now provide. In reality, however, we have so many requests from industry and other government agencies for increased accuracy, extended ranges, or measurement of new quantities, that we must carefully investigate all areas to determine where we can best use the talents of our staff, not only to provide immediate improvements but also in doing research to develop techniques of measurement on which new standards can be based. Our Radio Standards Laboratory, for example, has established a long-range planning committee and is setting up plans for the next five years or so. This is one step in the right direction, though more are needed.

In order to properly identify measurement needs and plan for the future, the Bureau must be able to obtain solid information and competent advice on the country's needs in all measurement areas. For this information we rely on a number of channels:

NBS has a contract with the National Academy of Sciences-National Research Council to assemble a number of Advisory Panels, composed of technical experts, to assist the various divisions of the Bureau in the formulation of long-range research plans. These panels meet periodically with the Bureau divisions they advise and closely examine the present--and projected--programs. The Director established a special Technical Advisory Committee on Calibration and Measurement Services in 1961 to facilitate liaison between the Bureau as a whole and industry. This Committee advises the Bureau concerning the needs of industry for measurement and calibration services, and suggests means by which the unique competence and facilities of the Bureau can best be used to provide a consistent system of calibration and measurement services. The present membership of the Committee (which met here in Boulder during this Conference) includes:

William G. Amey
Research Division
Leeds & Northrup Company

Ivan G. Easton
General Radio Company

L. B. Wilson
Sperry Gyroscope Company

Bruno Weinschel
Weinschel Engineering

Sheldon C. Richardson
General Electric Company

Charles E. Johnson
Aerospace Division
The Boeing Company

George Sonnemann
American Optical Company

Charles E. White
Avco Research & Advanced
Development Division

Joseph E. Aldrich
Ryan Aeronautical Company

Another input source is the joint NBS-Air Force Working Group established in 1958. This group has the objectives of promoting standardization of precision measurement equipment and standards produced by Air Force contractors, evaluating measurement requirements beyond present capabilities and initiating planning actions to meet these needs, and determining the types and accuracies of standards desirable at successive measurement echelons. Representatives of the NBS-AF Working Group have made several recent trips to contractors around the country in an attempt to get first-hand information concerning measurement requirements.

Another source of information was the Aerospace Industries Association survey of how well the measurement needs of the aircraft and missile industry were being met. The results of this survey, which graphically revealed serious gaps in some areas, led, as I mentioned before, to a special appropriation from Congress to make a start in several areas, notably radio measurements. A continuing series of measurement research conferences has grown out of this original survey, and has helped maintain a free interchange of information between NBS and manufacturers.

Of course, we look to the NCSL as a major source of information concerning both present and projected measurement needs. This organization is in a unique position to achieve a nationwide view of what measurement problems exist.

6. National Conference of Standards Laboratories

NCSL will become a valuable organization, justifying itself to the management of the member institutions as it accomplishes a series of useful tasks. I would like to suggest a few: A roster of standards laboratories, classified as to who they are, and what they can do, would be a most useful document, and one which the organization could produce much easier now than anyone else. Such statistics as how standards laboratories are organized (and perhaps how they fit into the corporate structure), special or unique services they offer, number and experience of personnel, dollar volume of calibrations performed annually, and others would help formulate a much clearer picture of standards and calibration activities in this country.

The NCSL can also facilitate the pooling of information on a variety of topics, including various practices on recalibration intervals, spot checking of instruments, use of one-point calibrations, and other items of widespread interest. The development of "recommended practices" on various aspects of any calibration laboratory's technical operations, and of calibration procedures for widely used instruments, is a most important area in which the NCSL can make unique and valuable contributions.

7. Accuracy Charts

In an effort to make NBS capabilities in various measurement areas more widely known, I have been urging each Bureau section involved in measurements to prepare a chart on which the accuracy of measurement (or calibration) is plotted over the range of the quantity measured. This has not been an easy task, and it has been difficult to agree upon some standard format for these presentations. However, progress is being made, and some recent first attempts in some areas are contained in other papers of this session.

A generalized accuracy chart is shown in figure 1. It is a log-log plot of decreasing uncertainty (upwards) against the magnitude of the quantity being measured. In general, the shape of the "curve" follows a somewhat predictable course. This stems from the fact that errors are often (a) independent of the magnitude of the quantity being measured in low ranges, (b) proportional to the quantity in intermediate ranges, and (c) increase as the square of some higher power at higher ranges. Thus, the curve tends to ascend at lower ranges (a), level off in a midrange (b), and fall at the higher magnitudes (c).

I recommend to all laboratories in the standards and calibration field that they plot similar charts depicting their own capabilities. On such charts they could show not only their "best" accuracy, but also the accuracy which they hope to provide to their users on the next lower echelon, and perhaps down to the ultimate user--for it is the accuracy requirements of the ultimate user in his practical measurements that justify all the superstructure of calibrating echelons, both in your labs and in NBS.

When any one echelon has a real need for accuracy (represented on the chart) within a factor of 2 or 3 of the next higher echelon, a measurement pinch may be developing. If the need of the top echelon of an industrial standards laboratory is *above* the NBS curve when plotted on the chart (and unfortunately this appears to be true in a number of rapidly developing measurement areas), this represents a failure on our part to anticipate and provide for a service in advance of the needs.

Such charts not only are useful in clearly setting forth a particular laboratory's measurement capabilities, but they can serve to pinpoint problem areas, and help make management aware of the need for, and the value of, improvements in a calibration program.

It should be noted that the various curves may be shown for any laboratory, for example, one representing its maximum capability for accurate *measurement*, another the accuracy of its best or special *calibrations*, another its usual calibration.

The plotting of such a chart brings to the forefront the basic questions of precision and accuracy. Two measurements may be compared with a precision considerably higher than the accuracy with which the measured values are known in terms of the unit. Other questions, for which it is hard to find uniform answers, are these: (1) How is the plotted uncertainty related to the scatter of observational errors?--Obviously curves of quite different levels would result from different conventions on this point. (2) What confidence levels are associated with the plotted values? (3) What is known, or estimated, about the systematic errors?

This is not the place for an extended discussion on these points--experts will probably be discussing them at length for some time before we all settle on uniform practices in this area. However, I would interject a note of caution for those who might be led too easily to conclude that a line on an accuracy chart represents "accuracy" of some particular instrument. Not so. Accuracy is associated with a procedure called a measurement, planned or executed by people. Precision instruments, plus careful calibration, plus careful use--all these enter into accuracy of measurement.

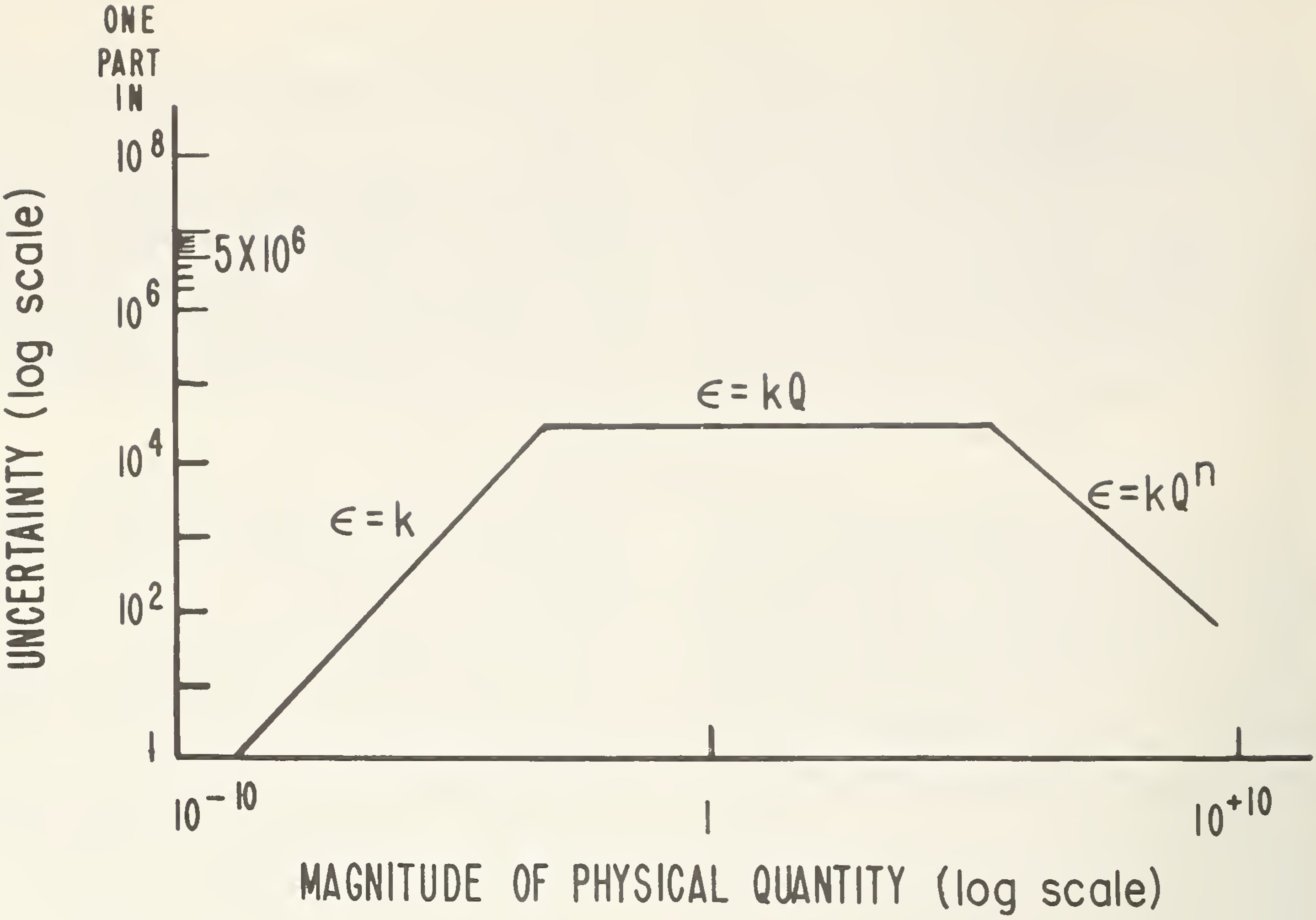


Figure 1. Generalized uncertainty chart.

Session 1. National Bureau of Standards Service to Industry

Paper 1.2. Length and Mass Calibration at the National Bureau of Standards

T. R. Young*

This paper provides a brief survey of the present length and mass measurement capability, developments underway, and plans for future work. Particular reference is made to present and future accuracy capabilities and requirements.

1. Length

Several formal changes have taken place in the field of length metrology in recent years--the consolidation of the various national definitions of the inch; the change of the primary standard of length from the International Meter Bar to a specific number of vacuum wavelengths of Krypton orange light. Another change has also taken place. This is the growth of activities and numbers of people involved in making length measurements to accuracies considered, not long ago, to be the province of the national standard laboratories. The state of the art of length measurement at the primary level, once considered to be of concern to a few, now is of interest to many; therefore, in this paper an attempt is made to present a state of the art chart based upon the experience of the Length Section and Engineering Metrology Section at the National Bureau of Standards.

Figure 1 illustrates this attempt. The chart is not complete nor is it fully self explanatory. It concerns only the accuracy of NBS calibrations of certain tools commonly used for linear length measurement; e.g., line standards, end standards including gage blocks, geodetic and surveying tapes. Generally, it does not indicate the accuracy to be achieved in the use of these tools. However, I believe that the chart is informative in that it indicates some advances, some present limitations and the probable direction of some future developments in length metrology.

The chart is plotted on a logarithimic scale to encompass the range and accuracies existing in length measurement. As an indication of accuracy, uncertainty taken as the three sigma limits of the respective calibration process, expressed in ratio to the total parts measured, is plotted against length, expressed in inches. Curves showing increasing accuracy for increase in length measured result from measurement processes that have errors independent of length measured as limitation to their accuracy. Such an error is the inability to index a measuring microscope exactly on the lines defining an interval of a line standard. Curves showing constant accuracy for all lengths measured are the result of measuring activities that have length-dependent errors as limitation to their accuracy. Uncertainty in the value of the coefficient of expansion of the material of the length standard can cause error of this type.

The curve for line standards is plotted over a length range from 20 μ in. to 48 in. Generally, reference standards in this length range can be compared with NBS standards on the Bureau's longitudinal comparator. The upper limit on the range is established by the capacity of the comparator. The lower limit on the range is established by the resolving power of the comparator's microscopes. For microscopes of the type used on this comparator it has been found from experience that observers differ in indexing the lines defining an interval by about 4 μ in. At present this provides the limitation to accuracy in the measurement of line standards in this size range.

The NBS 5-m is a line standard. The discontinuity of its curve from that of the line standard curve arises from the fact that it is measured in equipment not as easily controlled in regard to temperature, and is compared with a build-up of meter standards. Such build-ups in general create a loss of accuracy. In turn the 5-m bar is used to establish the location of the piers of the 50-m geodetic comparator. These piers support

---

*Chief, Length Section, Metrology Division, National Bureau of Standards, Washington, D. C.

micrometer microscopes that are used to calibrate the intervals of geodetic tapes. Such tapes are made of Invar to minimize expansion effect in field use. However, they are often subject to some instability. In fact the Bureau feels that only the most seasoned tapes are stable to 1 part in 500,000.

The geodetic comparator is also used to calibrate a NBS master steel tape which in turn is used to calibrate the steel tape bench. This bench consists of interlocking 12-ft sections of stainless steel joined together to form a graduated scale 200-ft long. Steel surveying tapes supported on the bench are normally calibrated by this scale to an accuracy of 0.001 ft. Optical tooling tapes having narrower graduation lines can, when the need is justified, be calibrated to an accuracy as high as one part in 400,000. However, this takes great care as the effective temperature of the 200-ft tape must be determined to an accuracy approaching 0.1 °C.

Transferring our attention to end standards brings us conveniently into the subject of using the wavelength of light as a reference length standard. Although a specific number of wavelengths of a radiation from krypton was adopted as defining the meter as recently as 1960, end standards, particularly gage blocks, have since the early twenties been measured in terms of wavelengths of light. Thus the techniques and uncertainties involved in light wave measurement are well known.

The chart indicates the advantages derived from the recent adoption. For a relatively small investment in equipment and training, anyone can apply the radiation from a Krypton 86 lamp to an optical instrument known as an interferometer. With this system a length scale can be obtained over 30 in. long with graduations approximately every 12μ in. This scale is accurate anywhere along its length to 1 part in 100,000,000 in vacuum and 1 part in 40,000,000 in ambient air. The accuracy of the latter is represented by the curve at the top of the chart. The advantage derived in increased accuracy over the line standard is obvious.

The accuracy of present methods of measurement of end standards illustrates the ability to measure material standards with the wavelength scale. The range of length extends from 0.01 to approximately 40 in. Over most of the range the limitation to accuracy arises from an inability to measure the displacement of the optical surface from the geometrical surface of the end standard to a certainty much better than 1μ in. For longer lengths additional uncertainty in determining the effective temperature of the end standard tends to level the accuracy curve.

The chart indicates the probable direction of future developments. Machines combining interferometric apparatus and photoelectric microscopes are already under development at NBS and other locations. These are designed to compare line standards against the wavelength scale. Further development needs to be made in instrumentation for indexing the rulings of line standards by application of complex microscopes designed for the ultraviolet and having long working distances and large numerical apertures. More accurate transfers between line standard and wavelength scales will then be possible. Investigations are also being made of new sources such as atomic beam, laser and microwave radiations to extend the length of the wavelength scale. If successful, more accurate measurements of longer length will be possible. To secure significant advantage in measurement of longer lengths of end and line standards against the wavelength scale it will be necessary to push the wavelength scale to higher accuracy. Measurement of longer lengths in a vacuum or the development of improved methods of sampling and determining the variables that affect the density of the ambient atmosphere will allow this to be accomplished. At the present time an air refractometer is being developed at the Bureau to measure directly the affect of ambient atmosphere upon the wavelength scale. The new tape standardization tunnel at the Gaithersburg facility has been designed with utilization of wavelength scales in mind. Methods of stability improvement of geodetic tapes are being investigated.

An investigation is well under way in the development of more accurate methods of measurement and manufacture of end standards known as gage blocks, particularly those in the size range 0.01 to 4 in. This project, from its inception, has benefitted from the support of a large number of private agencies. Major phases of this project are concerned with the development of measurement instrumentation and procedures necessary to achieve accuracy indicated by the broken line curve and the development of gage blocks with sufficient stability to warrant such accurate calibrations. Under this program prototype gage blocks have been developed with instability limited to 1 part in 10,000,000 per year and an interferometer that will measure optical length to this accuracy is near completion. To convert this optical length to practical length without loss of accuracy, microlengths such as thicknesses of wringing films will have to be measured to an accuracy approaching a molecular diameter. An application of an optical phenomenon known as frustrated total reflection is being investigated for this purpose. The curve shows the expected accuracy to be obtained in such measurements.

2. Mass

To conclude this talk I have three charts to show that indicate the state of the art in mass measurement. It is my understanding that the basic concept of using a balance to step down and to step up from the standard kilogram has served us very well in terms of supplying the accuracies of mass measurement that have been required. The advances that have occurred in the recent past have centered about the more efficient utilization of the general method. New designs in balances and in balance housings have reduced damping times and have minimized turbulence effects so that weighings that used to require hours and demand highly restricted test areas now may be performed in minutes in general-purpose measurement laboratories.

Figure 2 shows the state of the art in mass measurement in the same format as the length chart. The magnitude of uncertainty in ratio to the total parts measured is plotted against the mass measured in kilograms. As the ordinate here is the standard deviation, it should be multiplied by 3 to have the same significance that is shown on the length chart.

The various curves indicate the accuracy to be obtained by applying balances particularly designed for certain ranges of mass measurement; e.g., the quartz-fiber ultra-microbalance for milligram mass, the Rueprecht balance for the kilogram region, the Russell balance for the 2500-lb range. Mr. Lloyd who presented the next paper compiled this chart from information gained in measurement activities of the Mass Section at NBS.

Paul Pontius, Chief of the Mass Section, has compiled two other charts that are a valuable supplement to the state of the art chart for mass measurement. Two questions that always arise are:

1. What reliance can be normally associated with a set of reference weights?

and

2. What accuracy can be maintained in transferring from the reference weight to a given test weight?

Figures 3 and 4 provide answers to these questions based upon the measurement experience of the Mass Section. Standard deviation of a measurement, expressed in micrograms is plotted against the mass measured for a range up to 1 kg. The standard deviation computed for a number of NBS reference standards over a period of many years are shown by the triangular points. The two triangles at the 2-g level represent the latest single measurement and the measurement based on eleven redeterminations. The hexagonal point indicates the value obtained from all the redeterminations of the National Prototype Kilogram No. 20. The location of these points roughly indicates a reliance that can be associated with reference weights. The broken line curve indicates a "best" accuracy based solely on this reliance. The solid curve above these points is the basis for a control limit.

How well this accuracy in mass measurement can be maintained when measuring test weights is shown by the square and circular points in figure 4. For a balance suited to the mass measured, the standard deviations of single measurements are shown. This is representative of the accuracy that can be obtained in mass measurement by using available equipment and known weighing procedures.

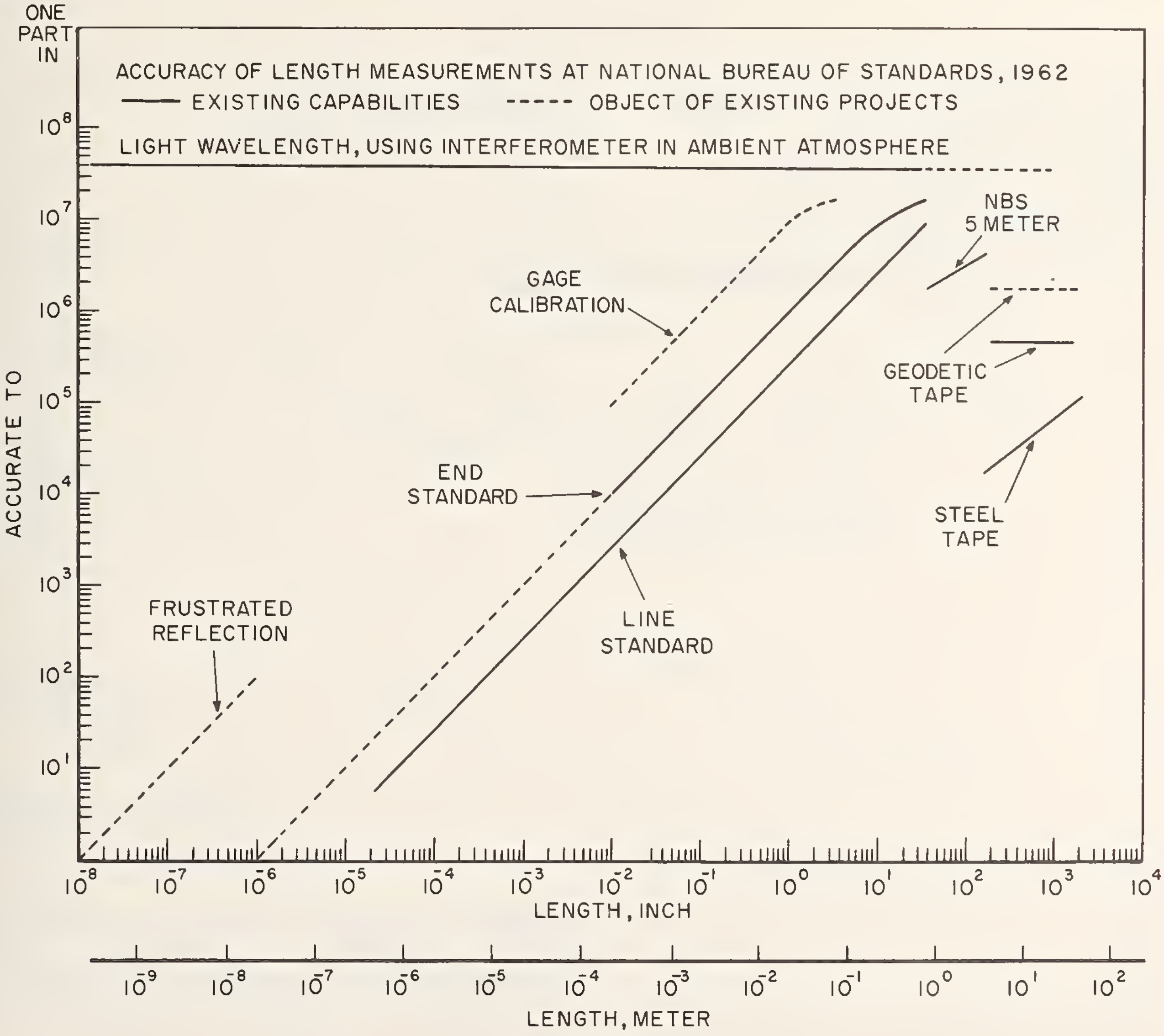
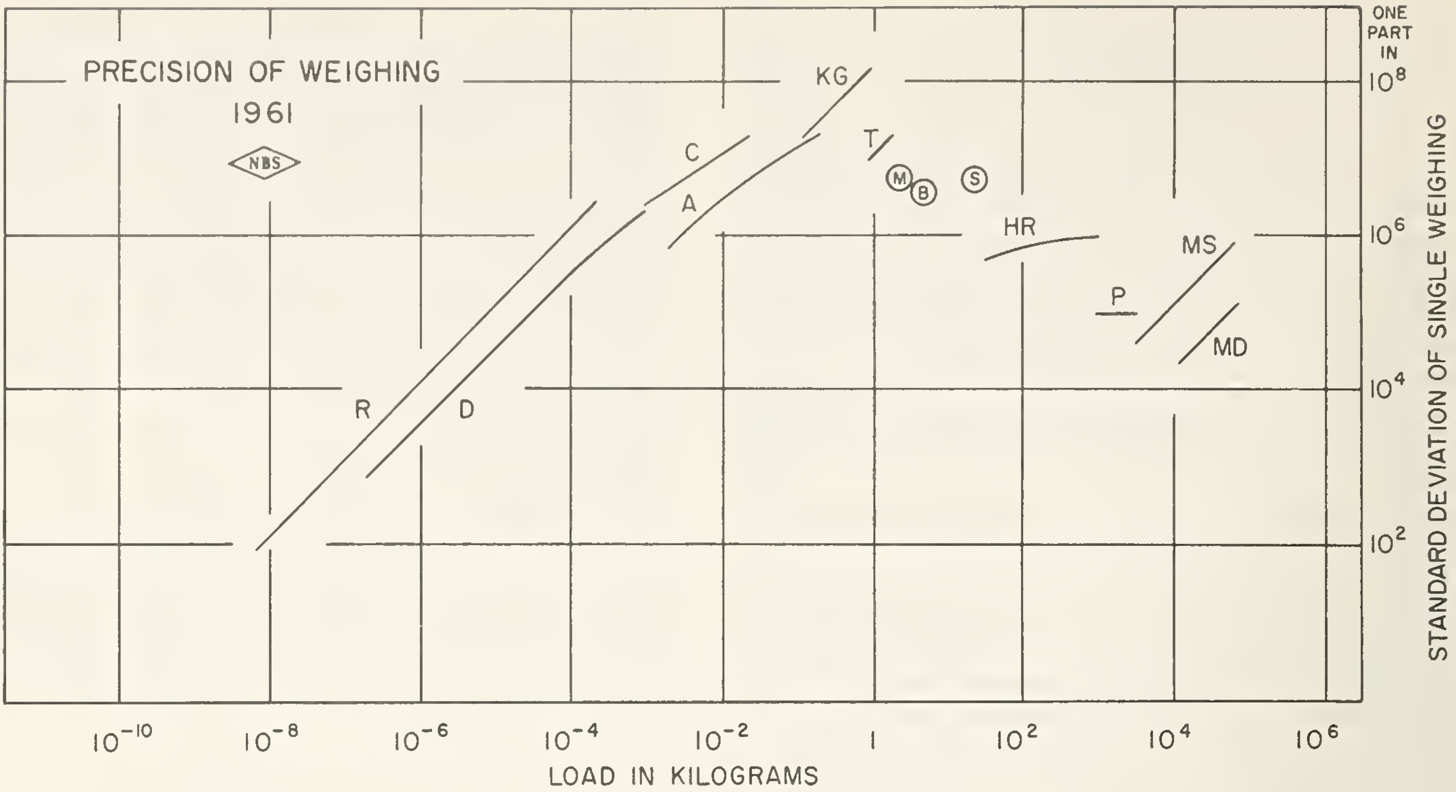


Figure 1. Accuracy of linear length measurement at the National Bureau of Standards. Uncertainty limiting the accuracy is represented at the three sigma level.



T Selected Equal-Arm 2kg Balance (NBS T-1)
M Quick-Weighing Single-Arm 6-Pound Balance
B Selected Equal-Arm 10kg Balance (NBS B-1)
S {Special 25kg blance (NBS S-1)
Quick-Weighing Single-Arm 50-Pound Balance
R Quartz - Fiber Ultra-Microbalance
D Special Assay Balance
C Corwin Balance

A Selected Equal-Arm 200g balance (NBS A-1)
KG Rueprecht 1 kg balance
HR Russell Balance, 2500 Pounds
P Platform Scale, 10,000 Pounds
MS Master Scale, 150,000 Pounds, Substitution Weighing
MD Master Scale, Direct Reading

Figure 2. Standard deviation of a single weighing achieved in mass measurement at the National Bureau of Standards using appropriate weighing devices.

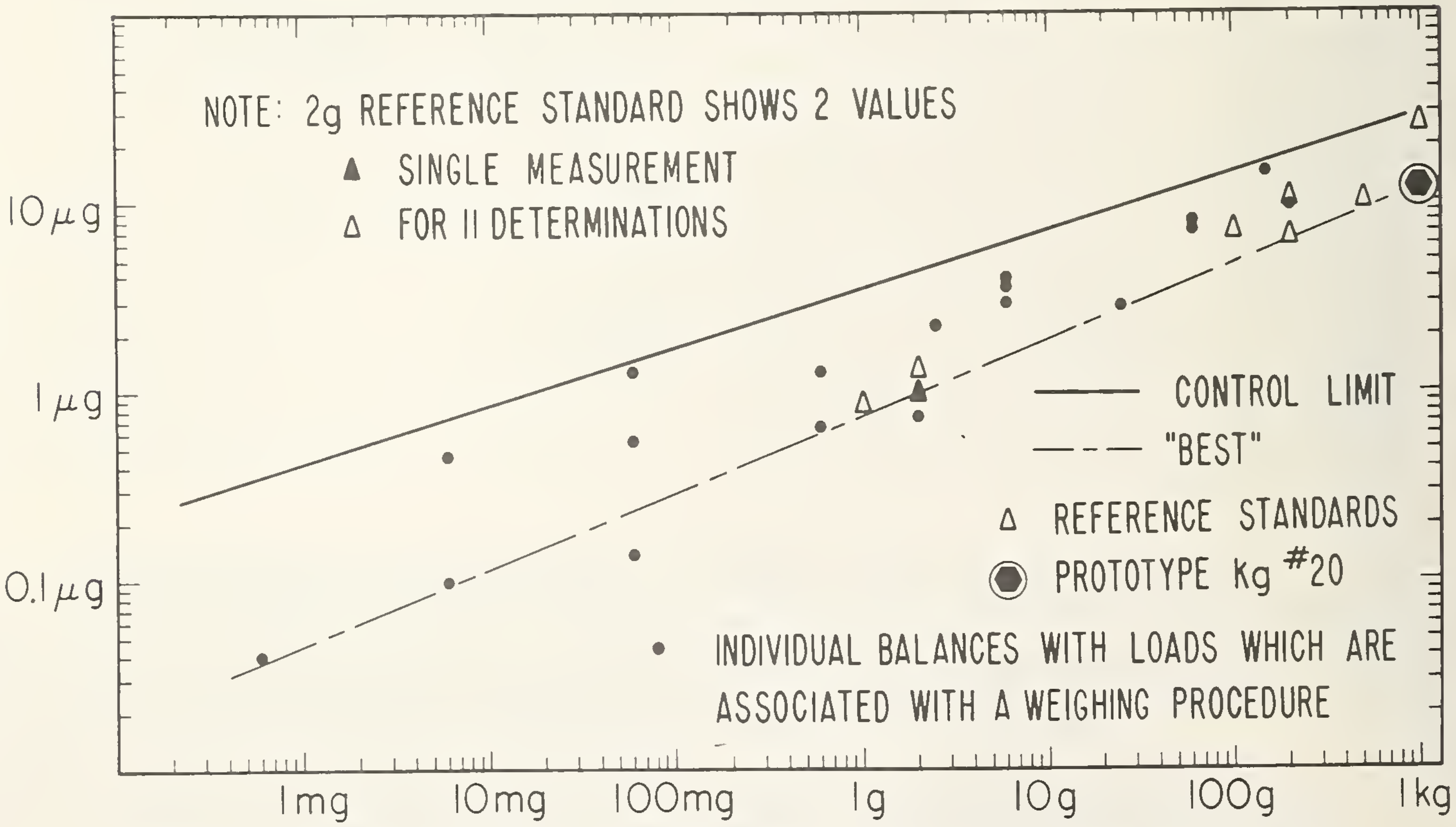


Figure 3. Standard deviation of a single weighing achieved at NBS compared to that for Prototype Kilogram #20, and for various NBS reference standards.

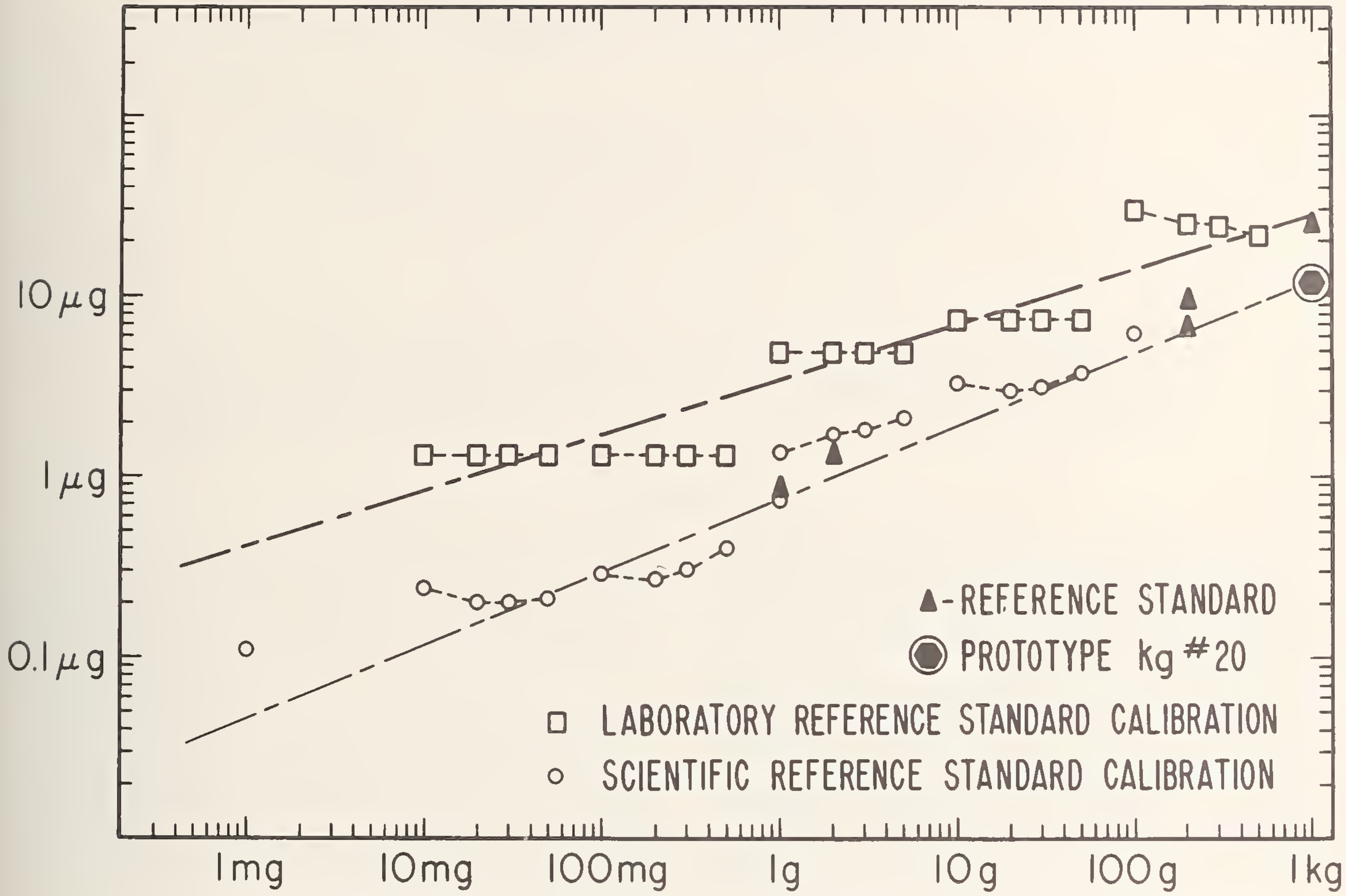


Figure 4. Standard deviation of the measured value achieved at NBS for laboratory and scientific type reference standards.

NBS Miscellaneous Publication 248
Proceedings of the 1962 Standards Laboratory Conference

Session 1. National Bureau of Standards Service to Industry

Paper 1.3. Pressure and Force Calibration at the National Bureau of Standards

E. C. Lloyd[1] and B. L. Wilson[2]

A summary of present NBS calibration capabilities and probable future requirements for pressure and force capabilities. Standards and techniques presently used and those under development for the measurement of pressures over a range of about 25 decades are described. The methods used for the measurement of forces up to several million pounds are described. Accuracies obtained are discussed.

1. Introduction

This paper deals with two measurement areas under the jurisdiction of the Mechanics Division--pressure and force. These are only two of the important areas for which the Division is responsible. The list also includes sound pressure and intensity, shock, vibration, strain, viscosity, and gas and liquid flow.

Nearly two years ago one of the large rocket engine manufacturers wrote to the Air Force "In any large rocket development program, in order to demonstrate the repeatability of total impulse specified for the propulsion system, the costs of a large number of tests are so high that the words 'cost' and 'accuracy' may be used interchangeably." Their analysis indicated that improved accuracy in thrust measurements could result in savings of many millions of dollars in the development of one large three-stage vehicle. It is thus understandable that the first special purpose laboratory to be placed in operation at the new Gaithersburg site will be the Engineering Mechanics Laboratory which will provide new facilities for calibrating force measuring devices by dead weights to 1 million pounds, or nine times the capacity of present NBS equipment. Pressure calibrations, also vital to rocket development and operation, will also be improved by providing new facilities at Gaithersburg.

2. Pressure Calibration

The relative importance of pressure calibrations is indicated by information from standards laboratories that in the mechanical area there are more instruments for pressure measurements than for any other quantity. A general view of the state of development of equipment for static pressure calibration may be obtained from a graphical presentation of measurement accuracy versus pressure. Figure 1 illustrates the accuracies now obtainable by a number of pressure measurement techniques.

This chart represents, from left to right, the range of pressures of more than 20 decades of present major interest in science and industry. This range extends from the high vacuum (10^{-14} mm of mercury at left of chart) that exists in parts of interstellar space, to pressures well over 1,000,000 lb/in.^{2} (10^8 mm of mercury, right end of chart) that are important in geology, in the study of properties of materials, and in some industrial processes. (These very high pressures are frequently given in kilobars--14,506 psi, approximately 1000 atm.)

Between these two extremes, 10^{-6} mm of mercury is the air pressure at an altitude of approximately 150 miles, 10^{-2} mm the pressure at approximately 50 miles, and 10^6 mm (10,000 lb/in.^{2}) the pressure at a depth of 20,000 ft under the sea, currently of new interest because of increased activity in oceanographic measurements.

The vertical scale at the right side of the chart shows the accuracy, expressed as the estimated

---

[1] Consultant, Mechanics Division, National Bureau of Standards, Washington, D. C.

[2] Chief, Mechanics Division, National Bureau of Standards, Washington, D. C.

uncertainty, with which we can presently measure these pressures.

The solid lines represent instruments now in use; the dashed lines show moderately optimistic hopes for the near future. Seven existing instruments are represented, including oil and air piston gages, four mercury columns, and a tilting diaphragm gage.

At pressures above 10 kb the measurements depend at present largely on calibration points determined by P. W. Bridgman and others; accuracies of these are estimated to be within about 1 percent at 15 kb and perhaps within 10 percent at 200 kb.

The shape of the curves of figure 1 merits some comment. In general any pressure-measuring instrument may have errors which are: (a) independent of the pressure being measured (e.g., errors due to zero instability, least count of reading means, capillary depression), (b) proportional to the pressure (e.g., proportional error in measurement of the height of a mercury column), and (c) errors relatively more important at high pressures such as the compressibility of mercury, or the elastic distortion of a piston and cylinder, for which the error is proportional to the square of the pressures.

At low pressures group (a) predominates; the intercept at accuracy 1/1 is drawn at the pressure equal to the estimated magnitude of these errors, and the curves start off with a 45° slope. At intermediate pressures the proportional errors (b) dominate and the curves are horizontal. (This is usually the region of greatest accuracy.) At high pressure the third group of errors (c) becomes important and the accuracy falls off.

These considerations apply not only to individual instruments but also to the state of the art as a whole. The envelope covering the individual curves has a maximum at a pressure of about a meter of mercury where, by great effort, it is possible to approach an accuracy of a part in a million with a mercury manometer. At higher pressures there is a definite loss in accuracy in measuring lengths greater than 1 m, and it becomes very difficult to maintain and to know the temperature of a mercury column more than a few meters in height. Also, the compressibility of mercury is about four parts in a thousand at 15,000 psi and is known to about 1 percent. Therefore accuracy of pressure measurement by use of a mercury column is limited at present to about $1/10^4$ at 30,000 psi. At the other end of the scale, the vibration level encountered in several of the National Laboratories makes it doubtful whether the level of a mercury surface can be found to an accuracy much better than $0.1\,\mu$.

We would like to mention briefly some of the work under way or planned at the National Bureau of Standards, having the aim of improved pressure calibration accuracy. For this purpose the 20 to 25 decade pressure range of interest can be broken into three ranges, in each of which the equipment and techniques tend to be significantly related. These are the range of very high pressures arbitrarily taken here as pressures above 25 kb, the vacuum range taken as pressures below $1\,\mu$ of mercury, and the intermediate range taken to be pressures between vacuum and high pressure.

In the high pressure range polymorphic transitions of various metals are useful as pressure reference points. In the last two or three years much effort has been expended in investigating the usefulness of some of these transitions, in particular the bismuth transitions above 25 kb, in several kinds of apparatus. Equipment of the multianvil type constructed at the National Bureau of Standards for work in this range is shown in figures 2 and 3. These photographs show the NBS six-anvil equipment used with a conventional hydraulic press to generate pressures of more than 100 kb. The pressure-transmitting media here is the well-known pyrophilite and a 1/2-in. cube of this material is compressed between the six tungsten carbide anvils shown.

As the hydraulic press load is applied the six tungsten carbide anvils move equally toward the center of the cube, guided by two conical steel rings shown. Teflon sheets lubricate the anvils in the rings, and also provide electrical insulation. At first the pyrophilite extrudes between the anvils; this forms a gasket that retains the remainder of the material, and allows pressures to be reached limited presently only by the strength of the anvils.

Below about 25 kb, liquids can be used as pressure transmitting media, and several fixed points are available. One of these is the freezing point of mercury at 0 °C, about 108,000 psi, and the National Bureau of Standards is presently making a new determination of this value. The equipment being used to measure the pressure is a controlled-clearance piston gage recently constructed at NBS.[1][3]. This is shown in cross section in figure 4. This instrument has a 0.080 in.-diam tungsten carbide piston in a steel cylinder, with jacket pressure applied to the outside of the cylinder to adjust the clearance to an optimum value. Extremely small leak-rates have been experienced. (Fall rate of less than 0.001 in. per hour at 113,000 psi.) The fluid used is 50 percent Varsol and 50 percent Octoil-S (di(2 ethylhexyl) sebacate or dioctyl sebacate). Eighteen 16 in.-diam weights totaling 600 lb are used for pressures up to 120,000 psi, and a 1,000 lb of weights at the maximum pressure of 200,000 psi. Operation at 120,000 psi has been very successful, and a sensitivity of one part per million has been achieved at 50,000 psi.

In the intermediate pressure range, NBS has recently constructed several instruments, both for absolute and differential pressures, for use below about two psi. One group of instruments in this range is a series of manometers using mercury, oil, and water, and with the liquid levels sensed by use of precision micrometers and point gages. The first of these instruments (4 in. mercury column) has shown a repeatability on the order of a few microns of mercury.

For pressures below a few inches of water we undertook, several years ago, the development of the tilting air-piston gage [2]. This instrument, illustrated in figure 5, is a piston gage in which the axis of the piston and cylinder is inclined to apply an appropriate component of the load, in this case only the piston, along the axis. In one experimental instrument designed for pressures up to 12 mm of mercury, a reproducibility of 1 part in 100,000 of the range was observed, and an absolute accuracy obtained of better than one part in

---

[3] Figures in brackets indicate the literature references at the end of this paper.

10,000. This type of instrument should be an important addition to the equipment of pressure standards laboratories.

In discussing the vacuum range, figure 6 shows the area of interest in terms of pressures and accuracies.

The National Bureau of Standards program in vacuum standards has so far concentrated its effort on the development of absolute techniques rather than on empirical methods of measurement.

Figure 6 shows part of the accuracy chart of figure 1, with a series of curves representing the estimated possible performance of several instruments under construction or proposed. Curves for an earlier 2-in. mercury manometer, and for a tilting-diaphragm gage are shown for comparison.

The next instrument represented, in the direction of higher vacuum, is the interferometer oil manometer. The Knudsen pressure division apparatus [3] works on the principle of a McLeod gage in reverse; preliminary results with this equipment show that it is usable at pressures approaching 10^{-7} mm. of mercury, but that the best accuracy is in the neighborhood of 1 percent, at 10^{-5} mm and higher. The two remaining curves refer to three versions of a vane-type instrument for direct measurement of force over area. One of these has been constructed, using a quartz fiber for application of the restoring force to a pair of vanes, and study of characteristics of this device will be under way soon.

Of the several experimental vacuum standards mentioned, a brief description of the interferometer oil manometer [4] may be of interest. This instrument is shown in schematic form in figure 7. The oil reservoir has an area in the center separated from the remaining free surface, so that the difference in the oil level inside and outside the partition depends on the pressure difference. Interference fringes developed between the lower surface of an optical flat and the oil surfaces are used to measure the difference in level. "Octoil-S" is distilled into the manometer after outgassing the system. Using the green mercury line (5461 A) one fringe is equivalent to 1.83×10^{-5} mm mercury. Estimation of one-fifth of a fringe yields a sensitivity of 3.6×10^{-6} mm mercury.

Space has permitted brief mention of only some of the work in pressure standards at NBS. Much of this work is covered in more detail in published papers or in papers in preparation, and we refer you to the bibliography [5, 6] for a listing.

3. Force Calibration

The accuracy and the ranges of present and planned force standards are shown in figure 8. In the present small dead-weight machines the uncertainty or maximum error in applied load is one part in 10,000 for loads from 10 to 10,100 lb. In the large dead-weight machine [7] the uncertainty is one part in 5,000 for loads from 2,000 to 111,000 lb. For the range from 100,000 lb to 1.5 million lb, using previously calibrated proving rings [8] in parallel, the uncertainty is one part in 1,000. For the range from 1.5 to 3 million lb the uncertainty is one part in 500. For the range from 3 to 10 million lb the uncertainty is one part in 300.

New dead-weight machines [9] to be installed at Gaithersburg will have capacities of 113,000 lb, 300,000 lb and 1,000,000 lb. These machines and existing machines will provide static forces known to within one part in 20,000 for the entire range from 10 to 1,000,000 lb. Above 1,000,000 lb it is believed that a new 12 million lb testing machine will permit calibration by means of multiple devices in parallel with uncertainty of one part in 5,000.

It is interesting to note that individual load cells of 6 million lb capacity have already been built and calibrated. Rocket thrusts of 10^8 lb are already being discussed. We wonder whether our new 12 million lb testing machine will be adequate by the time it is operational.

Figure 9 shows the weights of the 111,000-lb capacity dead-weight machine. These include ten 10,000 lb cast iron weights about 7 1/2 in. thick by 7 ft in diameter and nine cast steel 1,000 lb weights.

Figure 10 shows a diagrammatic sketch of the present 111,000-lb capacity dead-weight machine. The weights are located between the first and second floors, and the forces are applied to calibration devices between the second and third floors. The hydraulic jack located at the third floor level supports the upper frame which carries the upper tension and lower compression head. The lower frame (2,000 lb) is a part of the first load increment applied to the calibration device, shown here as a tension device. The force required to move the lower frame is applied through the device. The 10,000-lb weights are linked together so that each inch of upward motion of the frame transfers another weight to the device. The 1,000-lb weights are controlled separately. The area in which the device is calibrated is temperature controlled.

Calibration loads exceeding 111,000 lb but not exceeding 300,000 lb are applied to large devices by means of three calibrated 100,000 lb capacity proving rings as shown in figure 11. The three calibrated rings are placed in a vertical testing machine. A thick steel plate is placed over the upper bosses of the three rings. The ring to be calibrated is placed on a hardened block on the plate above the three rings. The four proving rings are read simultaneously while the load is held constant or increased very slowly. The force applied to the ring being calibrated is taken as the sum of the forces measured by the three previously calibrated rings.

The set-up in the 10 million pound testing machine for the calibration of devices up to 1.5 million pounds by means of five 300,000 lb rings is shown in figure 12.

The set-up for the calibration of a 6 million pound load cell by means of three 3 million pound capacity dynamometers is shown in figure 13.

A new 12 million pound capacity testing machine is being designed for the new Engineering Mechanics Laboratory at Gaithersburg to replace the present 10 million pound capacity machine, which is more than 50 years old. An artist's sketch of this machine is shown in figure 14.

Figure 15 shows the architect's illustration of the Engineering Mechanics Laboratory as it will appear when construction is completed. The building is about 97 ft high and the pits for the large machines extend as much as 26 ft below grade.

Figure 16 is a photograph taken August 1, 1962.

4. References

[1] Cross, J. L., Reduction of data for piston gage pressure measurements, NBS Mono. 65, (in press).

[2] Hutton, U. O., A tilting air-lubricated piston gage for pressures below one-half inch of mercury, J. Research NBS 63C (Eng. & Instr.), No. 1, 47-57 (July-Sept. 1959).

[3] Schuhmann, S., Study of Knudsen's method of pressure division as a means of calibrating vacuum gages, trans. 9th National Vacuum Sym., p. 463 (MacMillan Company, New York, N.Y., 1962).

[4] Thomas, A. M., Johnson, D. P., and Little, J. W., Design of an interferometer oil manometer for vacuum measurement, Trans. 9th National Vacuum Sym., p. 468 (MacMillan Company, New York, N.Y., 1962).

[5] Brombacher, W. G., Johnson, D. P., Cross, J. L., Mercury Barometers and Manometers, NBS Mono. 8 (May 1960)

[6] Brombacher, W. G., Bibliography and index on vacuum and low pressure measurement, NBS Mono. 35 (Nov. 1961).

[7] Wilson, B. L., Tate, D. R., and Borkowski, G., Dead-weight machines of 111,000- and 10,100-pound capacities, NBS Cir. C446, (June 1943).

[8] Wilson, B. L., Tate, D. R., Borkowski, G., Proving rings for calibrating testing machines, NBS Cir. C454, (Aug. 14, 1946).

[9] Million-Pound Dead-weight Machine Designed, NBS Tech. News Bull. 43, 12 (Dec. 1959).

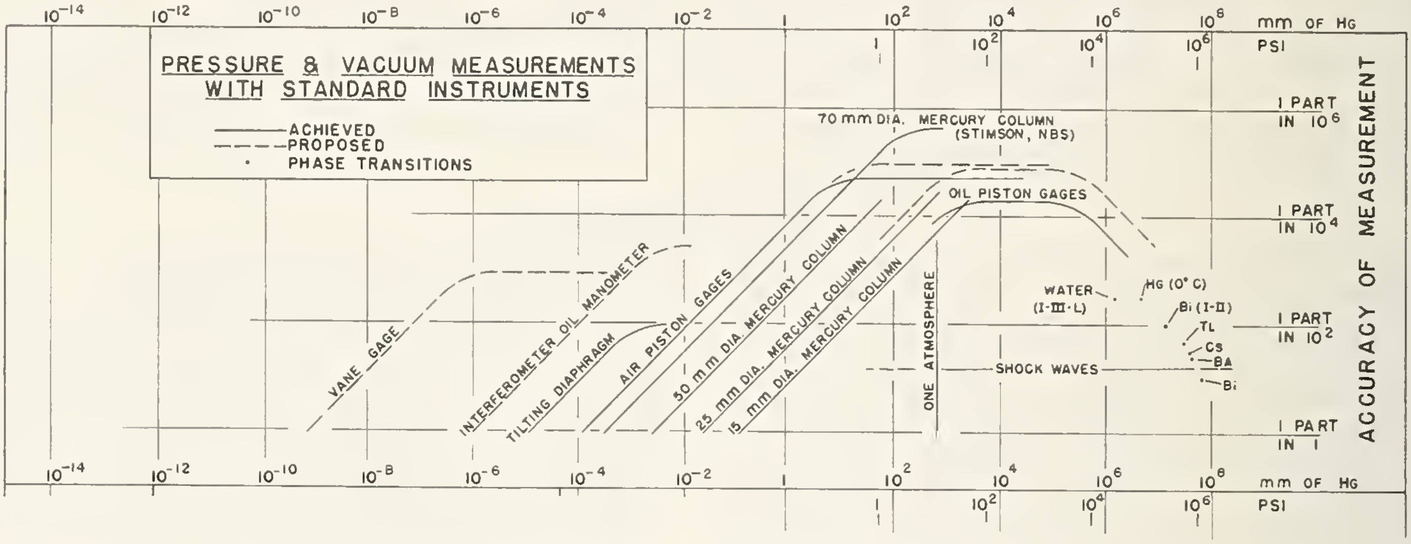


Figure 1. Pressure accuracy chart.

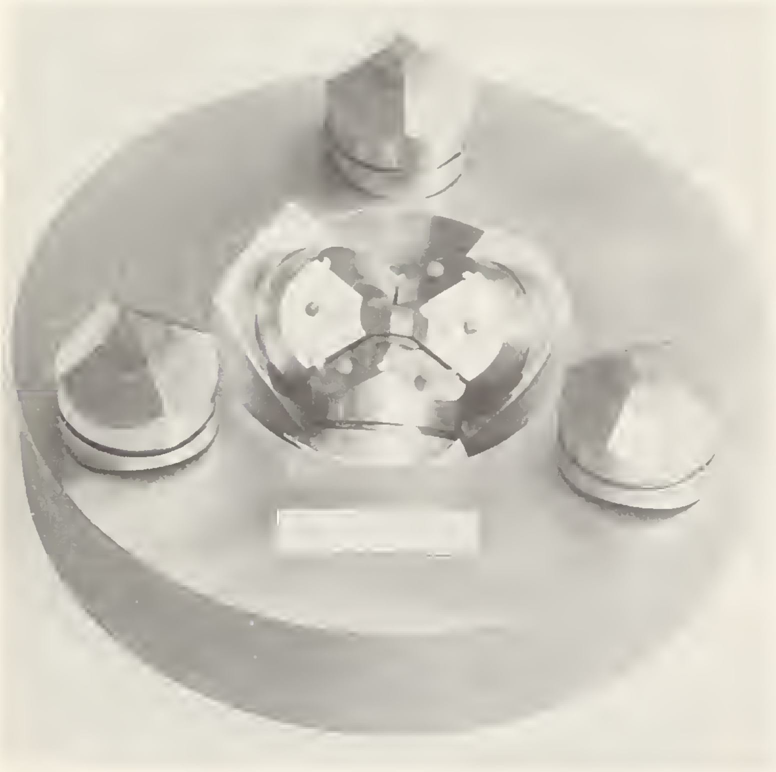


Figure 2. View of NBS six-anvil equipment for generation of very high pressures by compression of a one-half-inch cube of pyropholite. In use, the three anvils shown displayed on the ring are positioned against the upper three faces of the cube by a second identical ring.

Figure 3. Same as Figure 2, with five anvils in place.

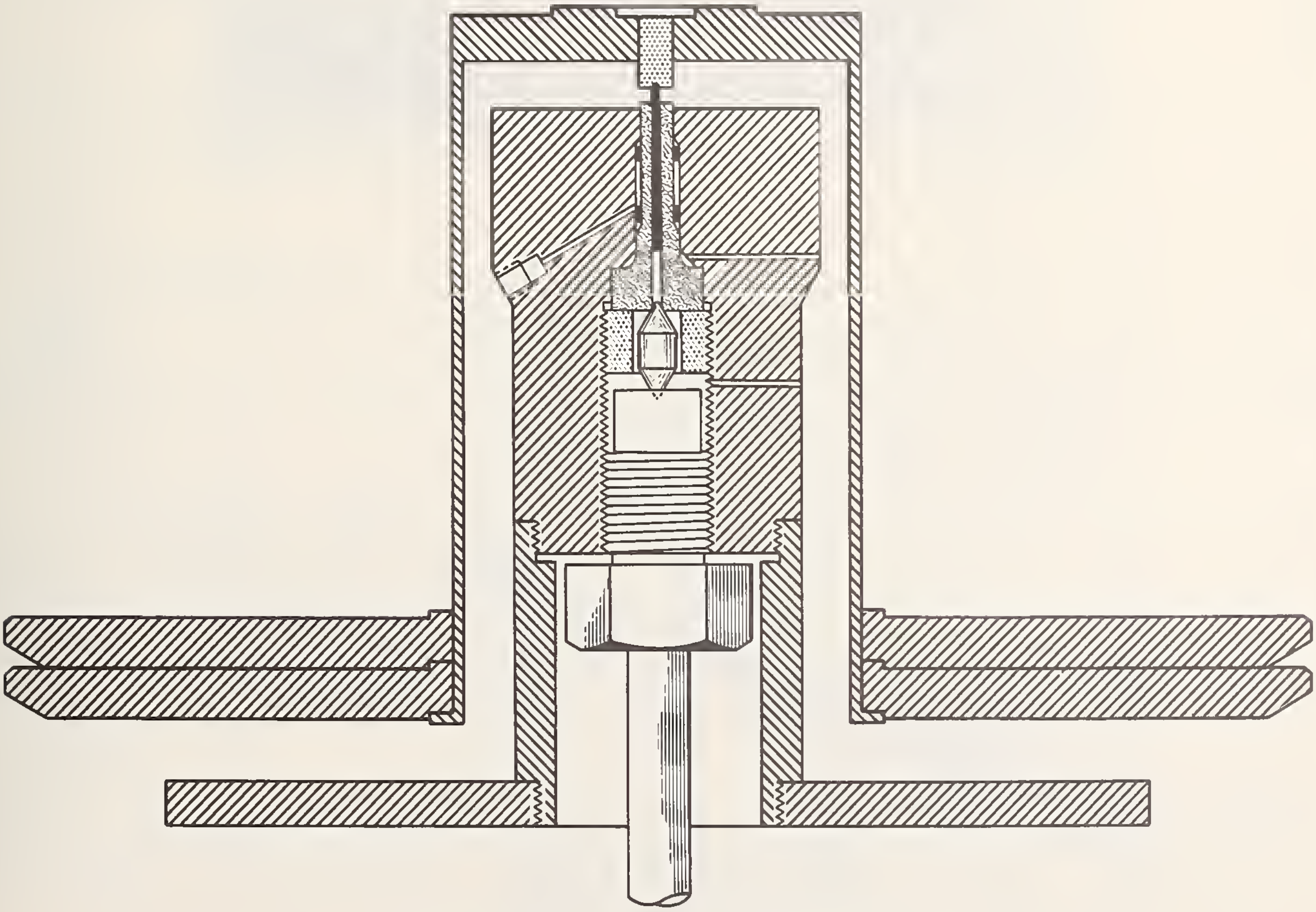


Figure 4. 14,000 kg/cm^2 piston gage.

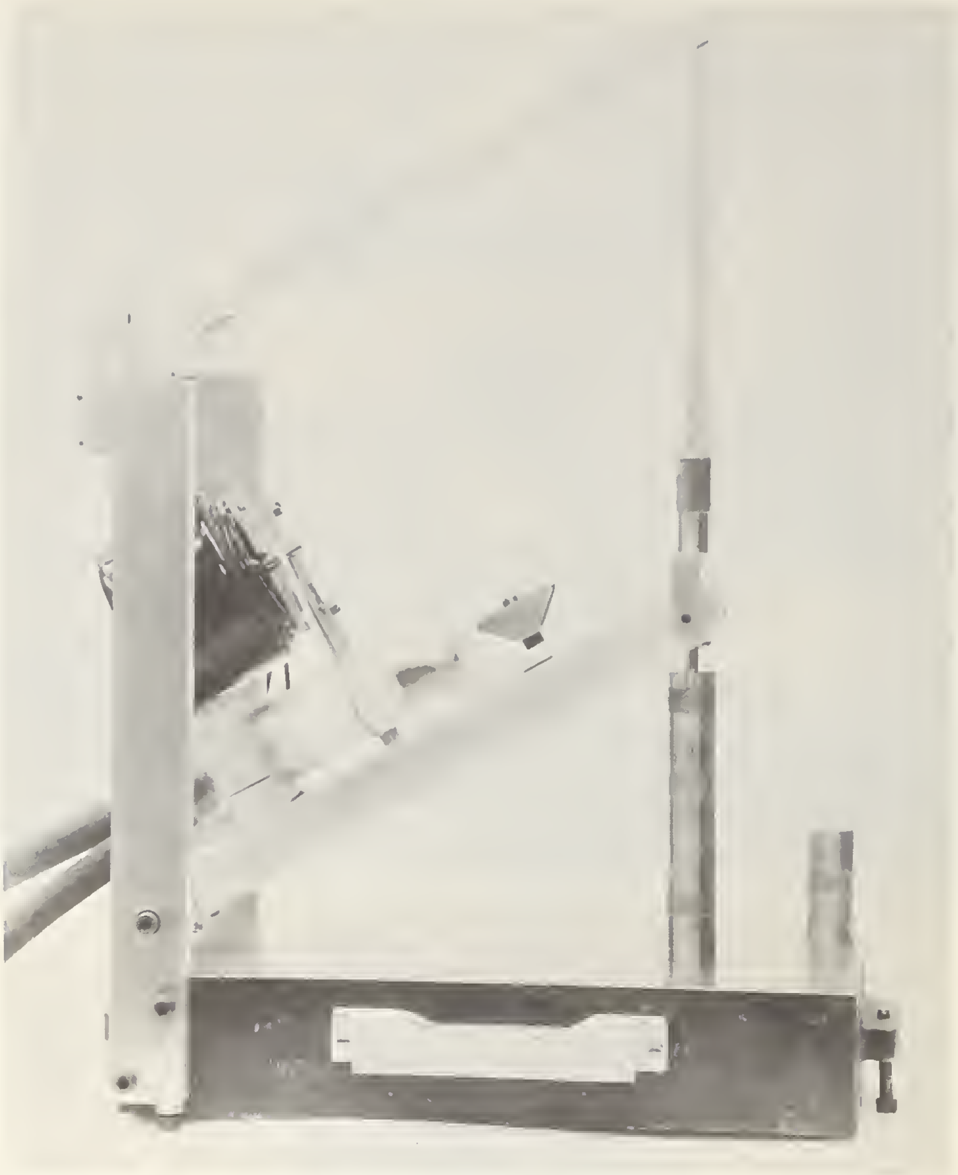


Figure 5. Tilting air piston gage.

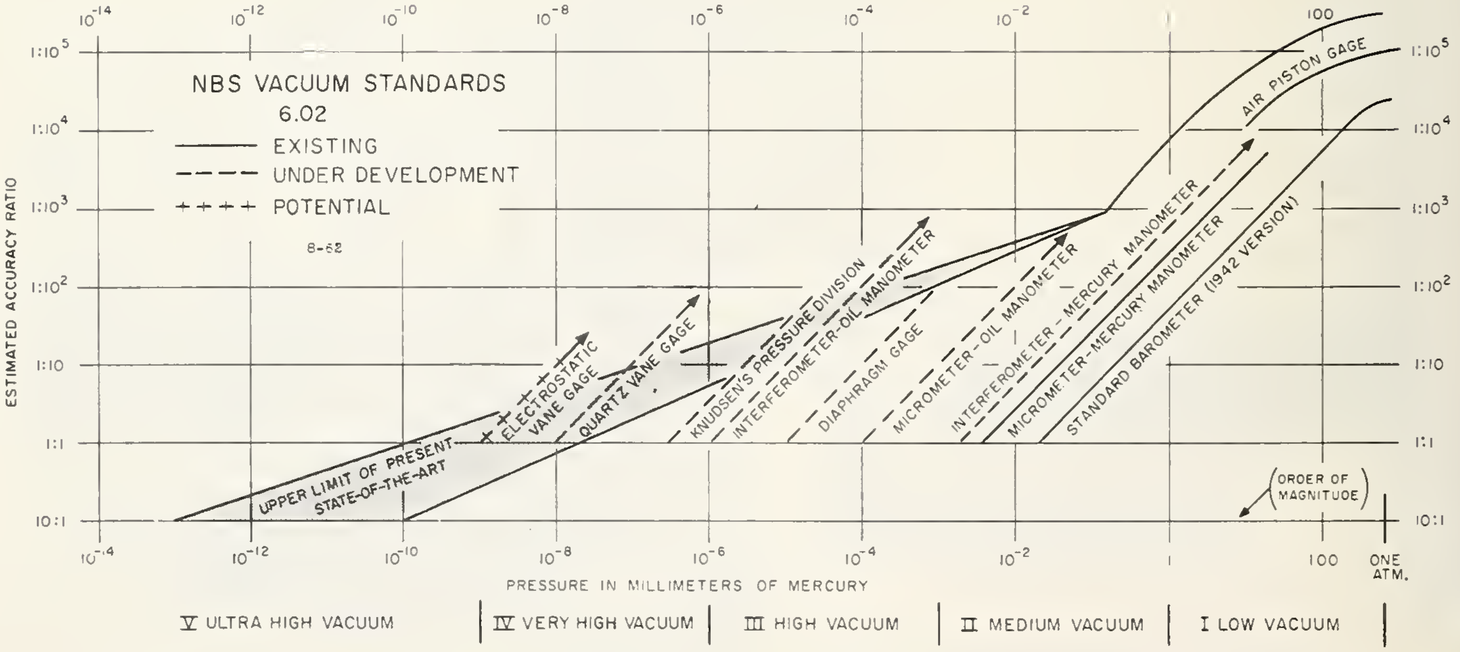


Figure 6. Vacuum accuracy chart.

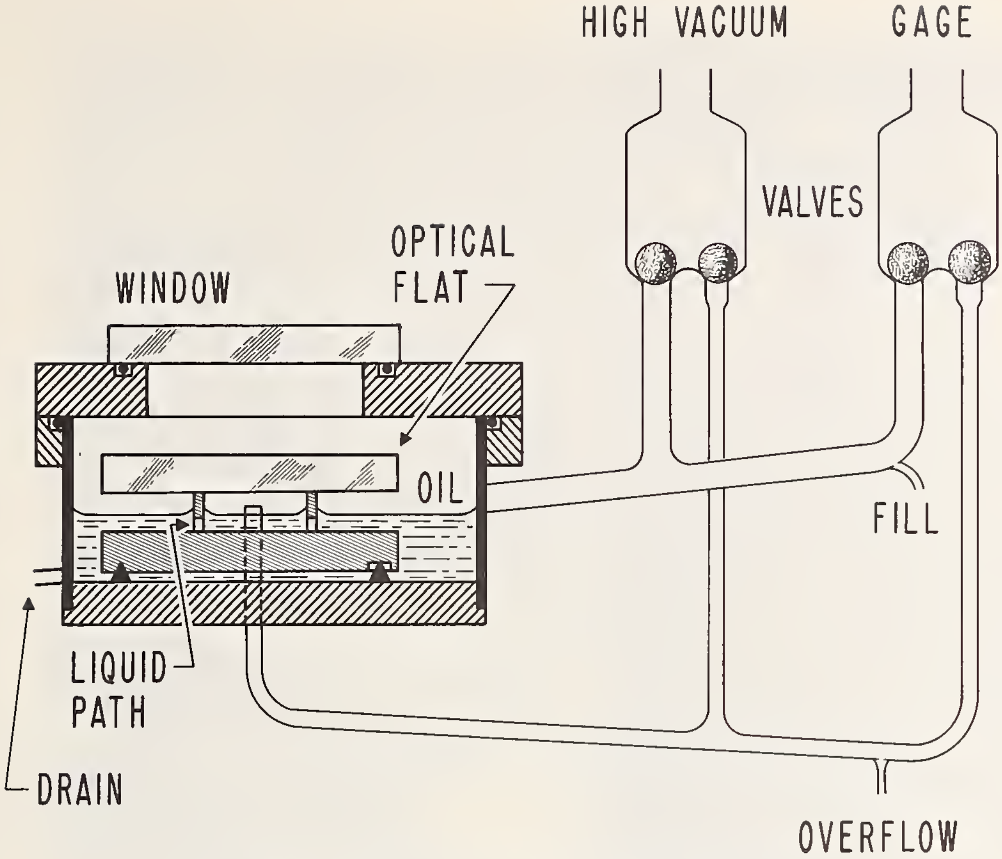


Figure 7. Interferometer oil manometer.

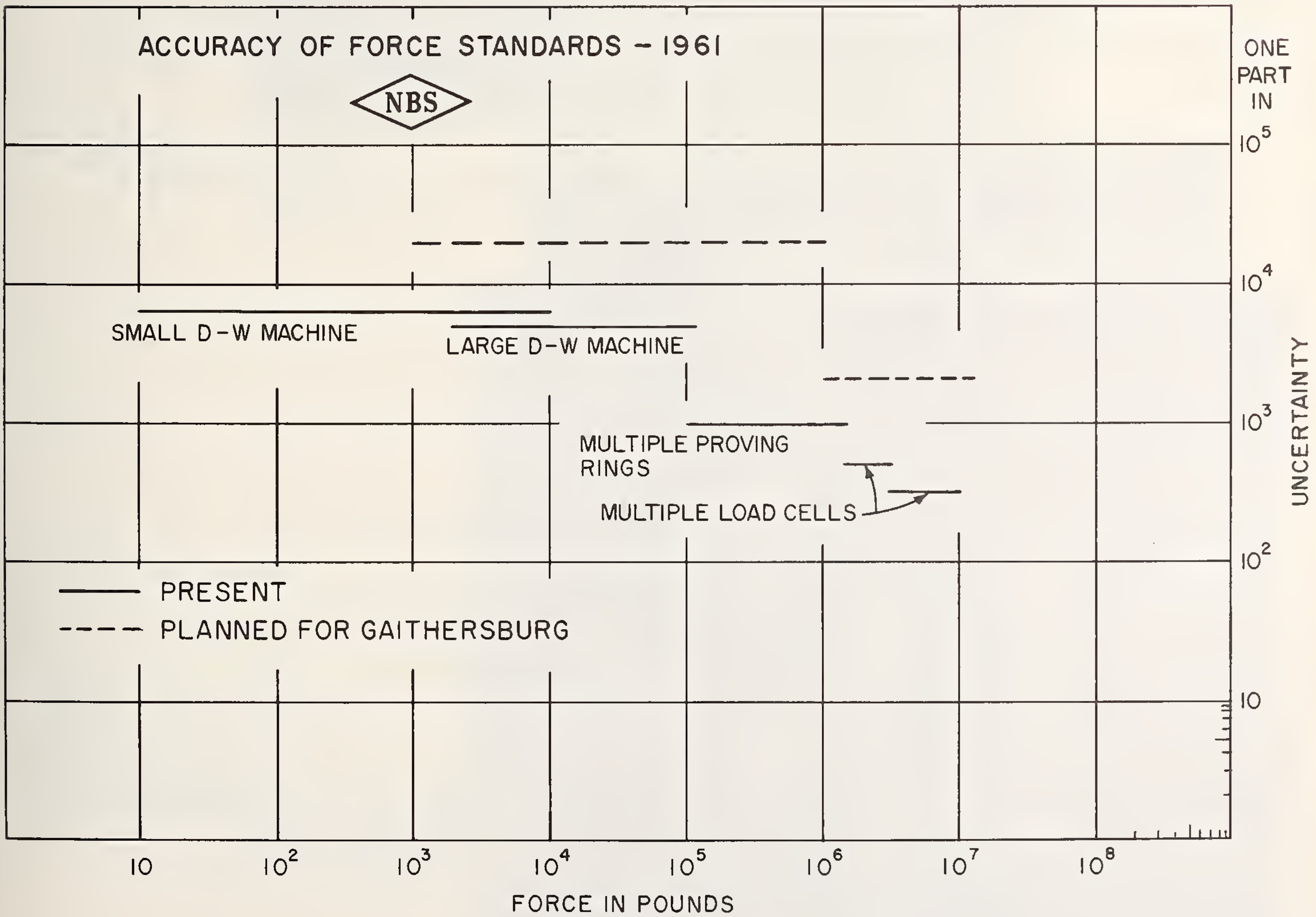


Figure 8. Force accuracy chart.

Figure 9. Weights and lower part of the 111,000-lb capacity dead-weight machine.

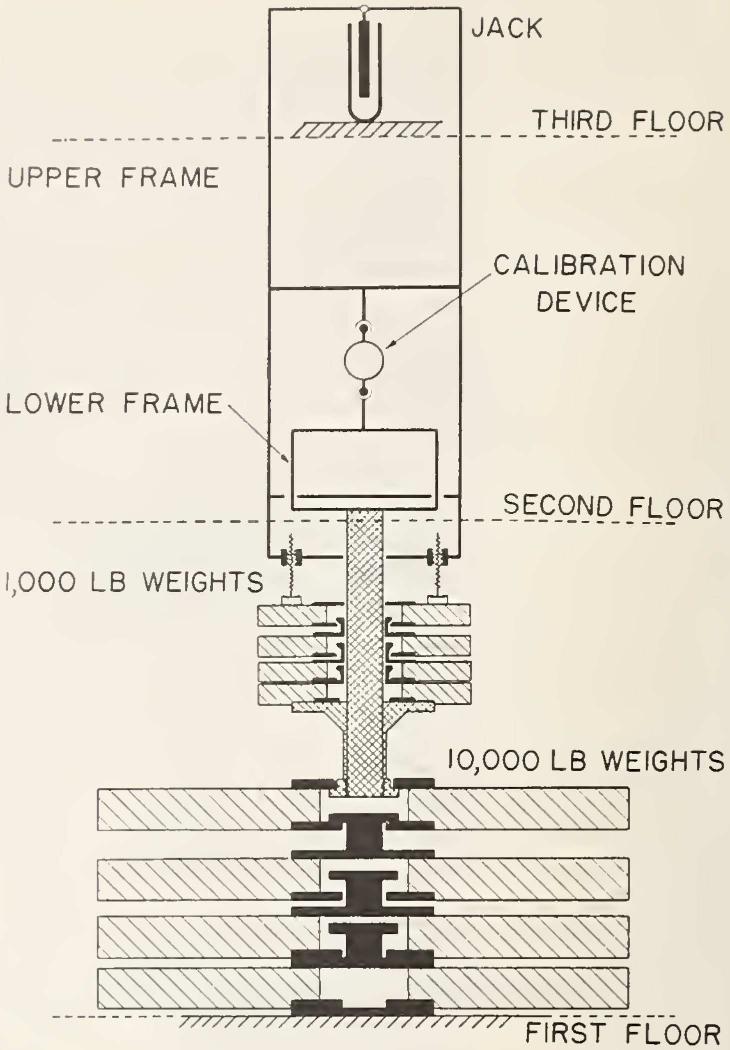


Figure 10. Diagrammatic sketch of the 111,000-lb capacity dead-weight machine.

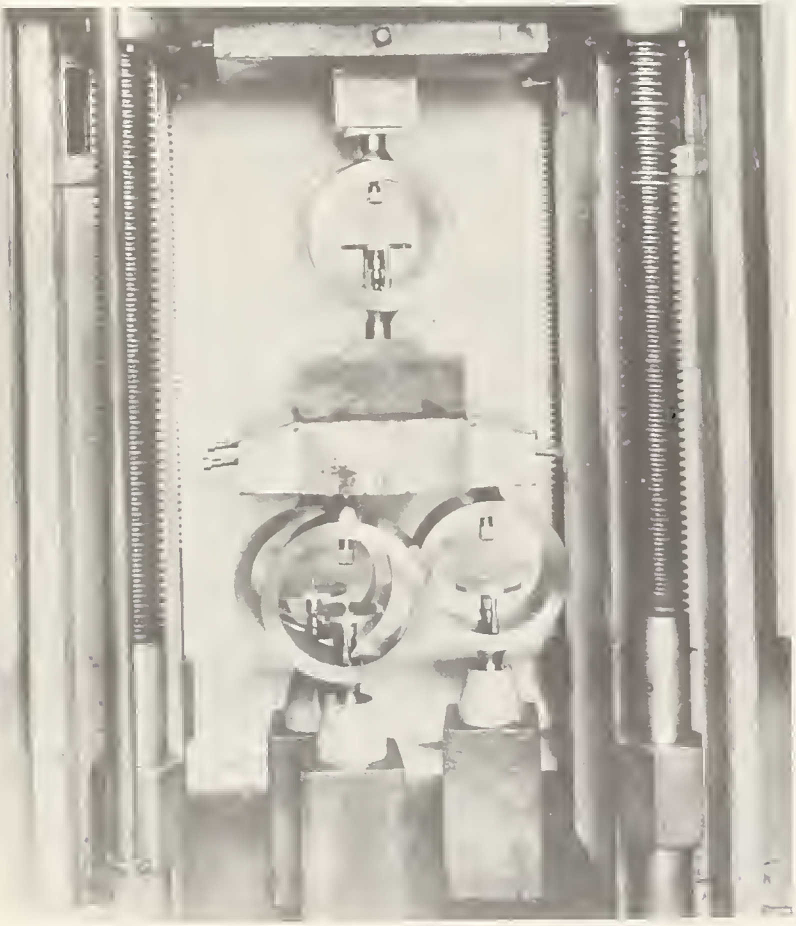


Figure 11. Calibration of a proving ring for loads exceeding 111,000-lb by means of three calibrated 100,000-lb capacity proving rings.

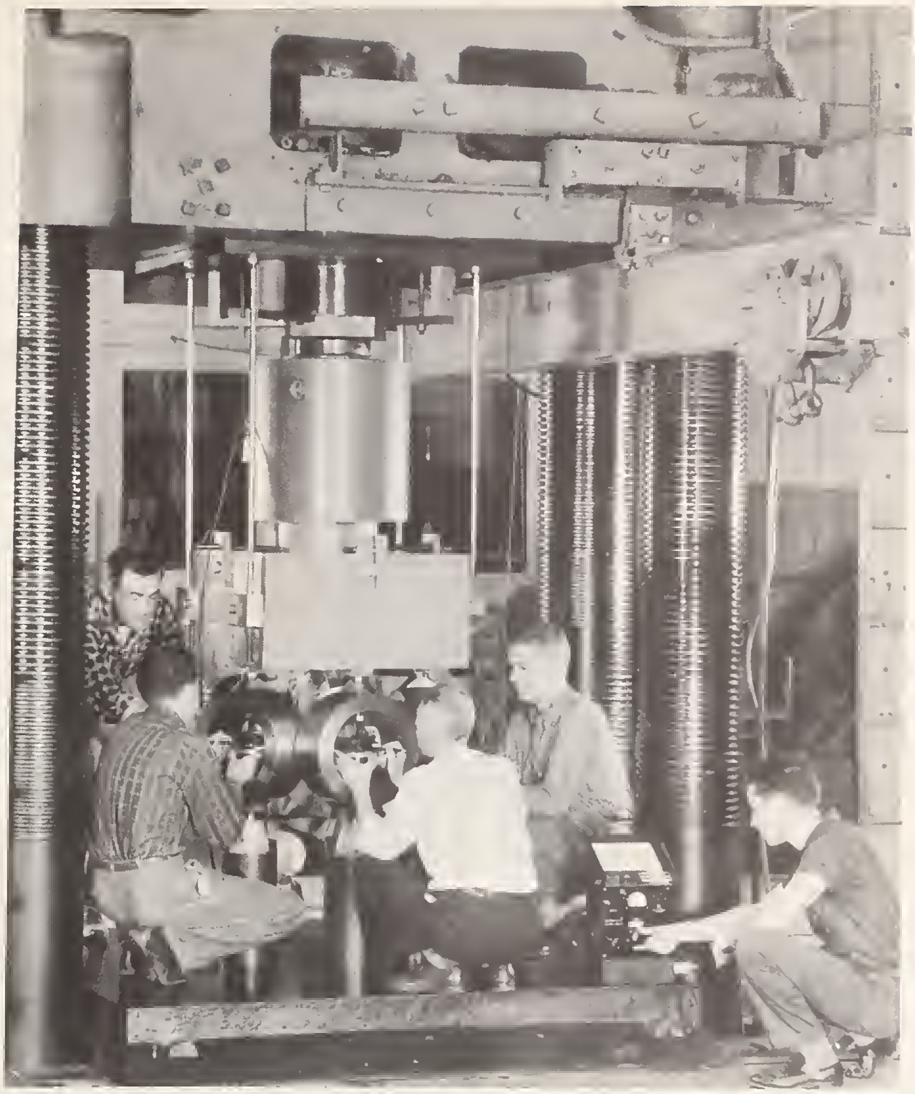
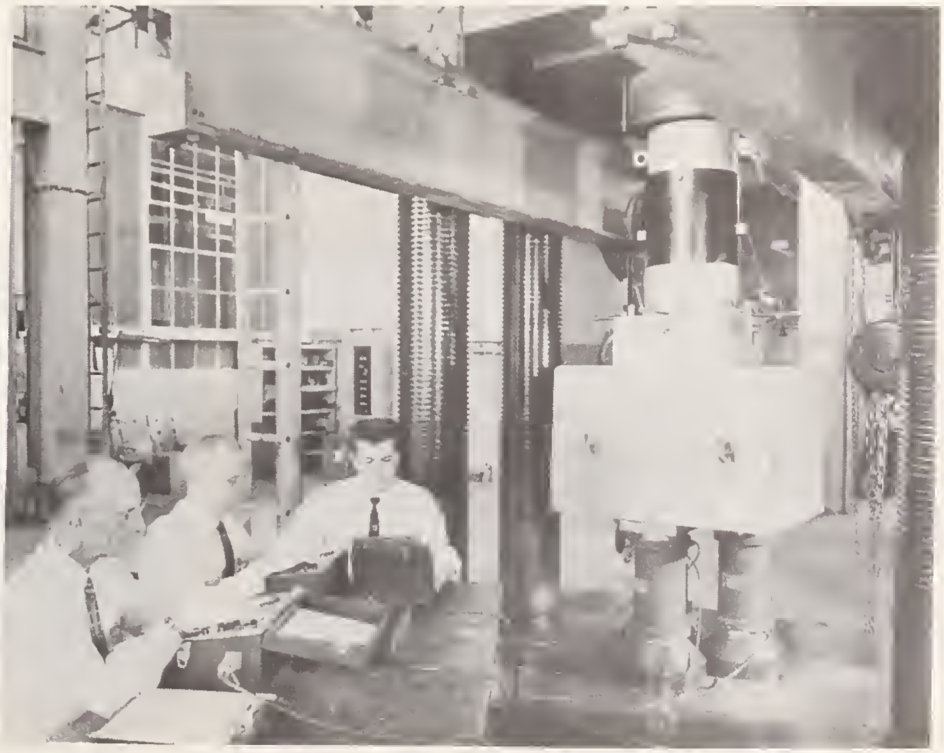
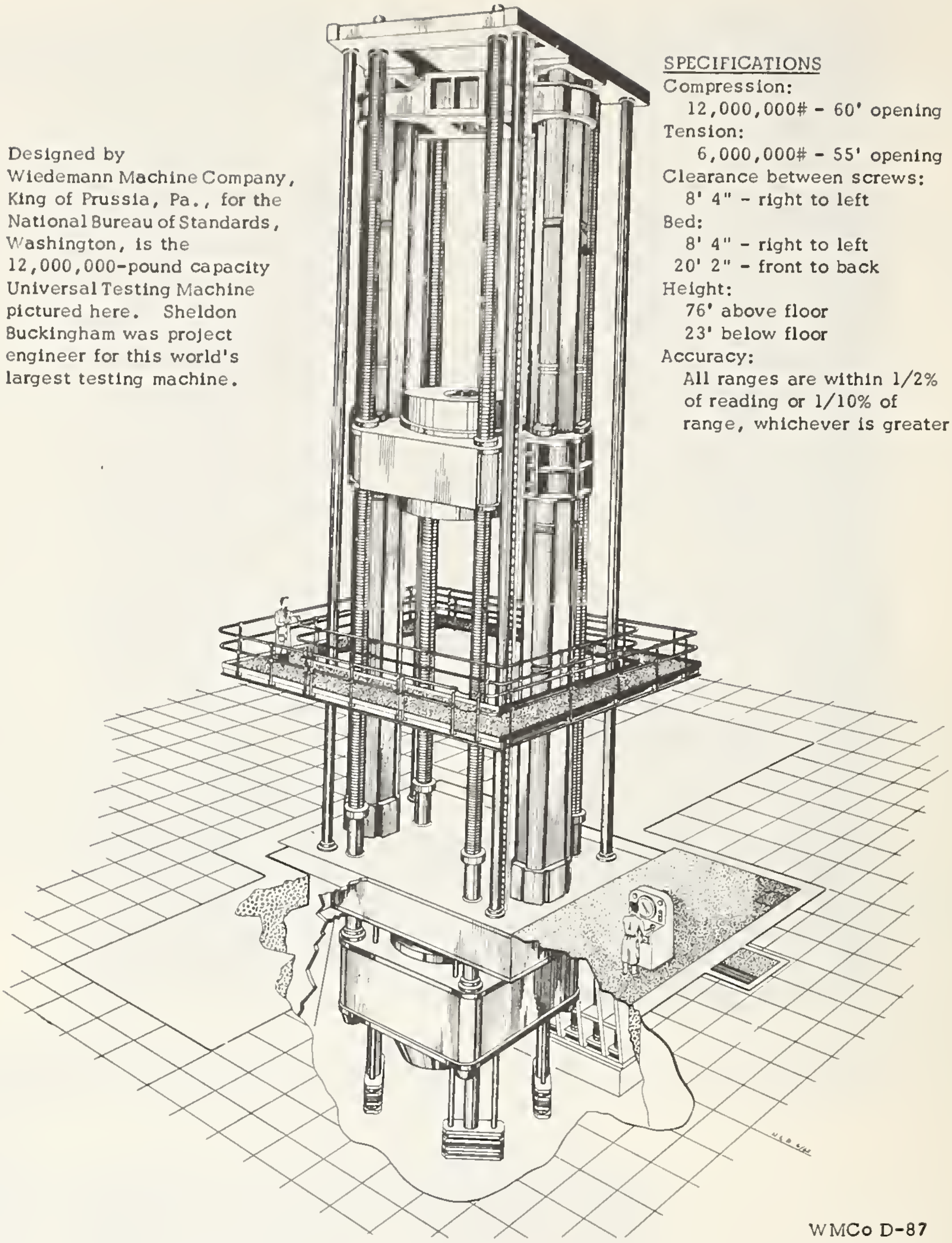


Figure 12. Calibration of a 1,500,000-lb capacity load cell by means of five calibrated 300,000-lb capacity proving rings.

Figure 13. Calibration of a 6,000,000-lb capacity load cell by means of three calibrated 3,000,000-lb capacity NBS dynamometers.





Designed by Wiedemann Machine Company, King of Prussia, Pa., for the National Bureau of Standards, Washington, is the 12,000,000-pound capacity Universal Testing Machine pictured here. Sheldon Buckingham was project engineer for this world's largest testing machine.

SPECIFICATIONS
Compression:
12,000,000# - 60' opening
Tension:
6,000,000# - 55' opening
Clearance between screws:
8' 4" - right to left
Bed:
8' 4" - right to left
20' 2" - front to back
Height:
76' above floor
23' below floor
Accuracy:
All ranges are within 1/2% of reading or 1/10% of range, whichever is greater.

WMCo D-87

Figure 14. Artists' sketch of the 12,000,000-lb capacity testing machine to be installed in the new Engineering Mechanics Laboratory at Gaithersburg.

Figure 15. Architects' illustration of the Engineering Mechanics Laboratory as it will appear when construction is completed.

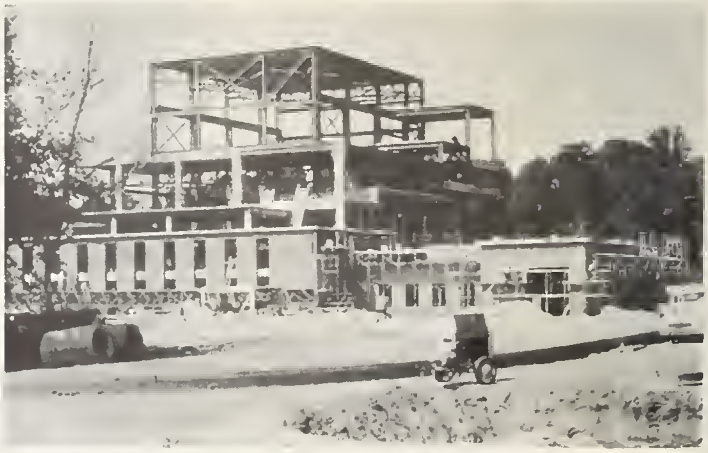


Figure 16. Photograph of the Engineering Mechanics Laboratory taken August 1, 1962.

Session 1. National Bureau of Standards Service to Industry

Paper 1.4. Calibration of Temperature Measuring Instruments at the National Bureau of Standards

James F. Swindells*

The National Bureau of Standards furnishes calibration services for the more precise types of temperature measuring instruments. The International Practical Temperature Scale and the NBS Provisional Scale of 1955 are maintained to serve as standards upon which these calibrations are based. Estimates have been made of the accuracy with which these scales are realized at the NBS, and, taking into account uncertainties introduced in the calibration process, limits of error have been assigned to specific calibration results. Work directed toward improving and extending these calibration services is a continuing activity.

1. Introduction

At the National Bureau of Standards, the International Practical Temperature Scale (IPTS) [1][1] is realized to serve as a common basis for defining temperatures in the United States. Below the lower limit of temperatures defined by the IPTS, a second scale, known as the NBS Provisional Scale of 1955 (NBS 1955 Scale), has been devised. These scales represent the National temperature standards for use throughout the scientific and technologic activities of the Country. Temperature measurements in scientific and industrial practice are related to these scales through temperature measuring instruments calibrated at the NBS in terms of the scales. To this end, calibration services are provided for both government agencies and private organizations or individuals. In general, only the more precise types of instruments, such as those suitable for use as laboratory standards, are accepted. The accuracies ultimately realized with such calibrated instruments in use depend upon several factors as follows: (1) The accuracy with which the IPTS is realized in the NBS laboratories; (2) the accuracy with which the instrument can be compared with the temperature scale; and (3) several use factors which involve the inherent limitations of the particular instrument and the soundness of the techniques which are employed in its use. In the following description of these calibration services, estimated limits are assigned to errors associated with the first two factors outlined above, but a consideration of use factors is beyond the scope of this paper. These limits of error are judgement-type limits based upon estimated magnitudes of the likely joint contributions of known sources of error and are used here as a measure of the degree of accuracy attained.

2. Temperature Scales Maintained at NBS

The International Practical Temperature Scale serves to define temperatures from -182.97 °Celsius (centigrade) up. This scale, which has been agreed upon by the 36 nations which subscribe to the actions of the General Conference on Weights and Measures, is defined by six reproducible temperatures to which values have been assigned. These "defining points" are the normal boiling points of oxygen at -182.97 °C, water at 100 °C, and sulfur at 444.6 °C (or the freezing point of zinc at 419.505 °C); the freezing points of silver at 960.8 °C, and gold at 1063 °C; and the triple point of water at 0.01 °C. The IPTS is further defined by specified instruments for interpolation between the fixed points together with

---

*Chief, Temperature Physics Section, Heat Division, National Bureau of Standards, Washington, D.C.

[1] Figures in brackets indicate the literature references at the end of this paper.

recommendations relating to their calibration at the appropriate fixed points. Thus the accuracy with which the scale may be realized varies in different temperature ranges throughout the scale and is in some cases dependent more upon the limitations of the specified instrument of interpolation than the accuracy with which the defining points may be realized. The standard platinum resistance thermometer is the specified instrument between -182.97 and 630.5 °C; from 630.5 to 1063 °C the platinum versus platinum-10 percent-rhodium thermocouple is specified. Above 1063 °C, a temperature on the scale is defined by the ratio of the spectral radiance of a blackbody at that temperature to the spectral radiance of a blackbody at the gold point (1063 °C). The instrument to be used for measuring these ratios is not specified, but a carefully designed optical pyrometer is currently used at the NBS.

In the temperature range covered by the platinum resistance thermometer, instrumentation in use at the NBS permits observations on a particular thermometer at a given time with very high precision. Limitations in the accuracy with which the defining points are realized, however, together with variations in the physical properties of the platinum resistors in different thermometers which affect the realization of the IPTS between defining points, restrict the translation of this precision into accuracy in the realization of the scale. Between 630.5 and 1063 °C, the accuracy with which the Scale is realized is limited principally by variations and instability of the physical properties of the standard platinum thermocouples. Above 1063 °C, in addition to uncertainties introduced by the optical pyrometer itself, one has an appreciable uncertainty resulting from observer error.

At the present time the lowest temperature defined by the IPTS is -182.97 °C (90.18 °Kelvin). The NBS has established a scale known as the NBS Provisional Scale of 1955 which serves to provide a temperature scale in the range 10 to 90.18 °K until such time as international agreement may be reached on a generally accepted scale. The NBS 1955 Scale is defined by the resistance-temperature relationships of a group of capsule-type platinum resistance thermometers which were originally calibrated by means of a gas thermometer. Occasional recalibrations have shown these thermometers to be sufficiently stable that no changes have been detected over the extended period that they have been in use.

Estimates of the limits of error in realizing temperatures on the two scales at the NBS are given in table 1. In the case of the NBS 1955 Scale, the precision of the resistance thermometers comprising the group used to maintain the scale is more than an order of magnitude better than the accuracy of the gas thermometry against which they were originally calibrated and against which they have been checked for stability at later times. With this situation it cannot be absolutely established that these thermometers are remaining stable to better than the limit of error assigned to the gas thermometry (ca 0.01 deg). Intercomparisons of thermometers within the group, however, indicate that they have changed much less than this, unless the unreasonable assumption is made that the thermometers have all changed by very nearly the same amount. The limits of error assigned to the realization of this scale are estimates based upon this reasoning. In the temperature range covered by the IPTS, comparisons of the scale as maintained in the national laboratories of several different countries serve as the basis for a reliable estimate of the limits of error throughout the scale as realized at the NBS. Accuracy estimates for both scales as realized at the NBS are represented by solid lines in figure 1. In the figure, temperatures on the International Practical

Table 1. Temperature scales used as a basis for NBS calibration services

Temperature		Scale	Thermometer used to realize scale	Estimated limits of error in realizing scales
°K	°C			
				deg
11.	-262.15	NBS 1955	Resistance therm.	0.004
20.	-253.15	...do...	........do.......	.001
90.18	-182.97	...do...	........do.......	.001
90.18	-182.97	IPTS	........do.......	.005
273.16	+0.01	...do...	........do.......	.0002
373.15	100.	...do...	........do.......	.0005
717.75	444.6	...do...	........do.......	.002
903.65	630.5	...do...	........do.......	.01
903.65	630.5	...do...	Thermocouple	.2
1233.95	960.8	...do...	........do.......	.2
1336.15	1063.	...do...	........do.......	.2
1336.15	1063.	...do...	Optical pyrometer	.4
2273.	2000.	...do...	........do.......	2.
4273.	4000.	...do...	........do.......	10.

Kelvin Scale[2] are plotted as abscissas on a logarithmic scale. Also shown, is a scale in degrees Celsius.

Plotted as ordinates, also on a logarithmic scale, are accuracies defined as estimated limits of error expressed as fractions of the Kelvin temperature. These fractions are reduced to 1 part in the appropriate power of ten. For example, a limit of error of 0.4 deg in a measurement of the temperature, 400 °K would be plotted as accurate to within 1 part in 10^3.

The large discrepancy apparent between the accuracies of realization of the oxygen point (90.18 °K) with respect to the two scales is an inherent consequence of the way in which the two scales are defined at this common temperature. In the case of the NBS 1955 Scale, the temperature is defined in terms of the mean of readings of a group of very stable thermometers, and this can be done quite accurately. The accuracy assigned to the realization of the oxygen point as a defining point on the IPTS, on the other hand, involves not only the reading of a resistance thermometer, but also the uncertainties inherent in the realization of the boiling point of pure oxygen. On this basis, less accuracy is assigned to the realization of this temperature with respect to the IPTS. The other abrupt changes in accuracy occurring at 903 °K(630 °C) and 1336 °K(1063 °C) are the result of progression from one instrument of interpolation to another as prescribed for the realization of the IPTS.

3. Accuracies Attainable with Calibrated Instruments

The essential feature of any temperature measurement (and sometimes a difficultly resolved source of error) is that the temperature sensor be at the temperature of the body or medium whose temperature is to be measured. But assuming no error from this source, whether or not the accuracies attained in the realization of the temperature scales maintained at the NBS can be transferred to an actual temperature measurement through use of a calibrated thermometer will depend principally upon the limitations of the thermometer itself. Usually, uncertainties introduced by the calibration process will not be significant compared with the sources of error indigenous to the thermometer being calibrated. The principal types of temperature measuring instruments regularly calibrated at the NBS are discussed briefly below.[3] Estimates of accuracies based upon assigned limits of error for selected calibration services are plotted as dashed lines in figure 1. These limits of error are largely judgement-type limits based upon estimated magnitudes of the known sources of error. In nearly all cases, there are insufficient data of a kind which will permit statistical analysis of the contribution of a particular potential source of error. Detailed descriptions of the calibration procedures and discussions of the limits of error assigned to calibration results are given elsewhere. [2 to 6]

In discussing the calibration services below it will be more readily meaningful to refer to limits of error in degrees, rather than accuracies, and to express temperatures on the NBS 1955 Scale in degrees Kelvin and temperatures on the IPTS in degrees Celsius. These limits of error represent minimums which can be assigned to the types of sensors discussed here.

Resistance thermometers will normally be calibrated only if they may reasonably be expected to meet the requirements as a standard on the IPTS. In general, this requires a 4-lead resistor of very pure platinum, mounted in a strain-free manner, and hermetically sealed in a protecting tube. Certain so-called capsule-type resistance thermometers intended for use at low temperatures meet these requirements and are calibrated down to about 11 °K. Limits of error attainable with the capsule-type thermometers relative to the NBS 1955 Scale vary from about 0.004 deg at 11 °K to 0.001 deg at temperatures between 20 and 90 °K. With standard thermometers calibrated in terms of the IPTS the limit of error decreases from about 0.005 deg at -183 °C to close to 0.0002 deg at the temperature of the triple point of water (+0.01 °C) and then gradually increases to about 0.001 deg at 200 °C and 0.01 deg at 630 °C, which is the upper limit of the resistance thermometer range. Employing good techniques these limits of error can be applied to temperature measurements made with this type of resistance thermometer.

Thermocouples of platinum versus platinum-10 percent-rhodium conforming to the IPTS requirements of a standard thermocouple may be given a "primary" calibration which consists of electromotive force (emf) determinations at the zinc point (419.5 °C), the antimony point (630.5 °C), the silver point (960.8 °C), and the gold point 1063 °C) with an estimated uncertainty equivalent to 0.2 deg at these temperatures. Interpolated values obtained by means of specified equations have assigned limits of error of 0.3 deg, while the limits gradually increase to 2 deg for values extrapolated to 1450 °C. An alternative service involving less accurate calibration is offered for both platinum versus platinum-10 percent-rhodium and platinum versus platinum-13 percent-rhodium thermocouples. Under this service, the test thermocouple is calibrated by comparison with a standard thermocouple at a number of selected temperatures between 0 ° and 1100 °C with the calibration extended to 1450 °C by extrapolation. Limits of error are 0.5 deg up to 1100 °C and gradually increase to 2 deg at 1450 °C.

Base metal thermocouples are also calibrated over the range 0 ° to 1100 °C by comparison with a standard platinum thermocouple in an electrically heated tube furnace. A 1-deg limit of error is assigned to this calibration. More accurate calibrations are offered for base metal thermocouples in the range -196 °C to +500 °C using stirred liquid comparison baths and resistance thermometer standards. Limits of error are 0.1 deg up to 300 °C and 0.2 deg from 300 to 500 °C.

---

[2] Values on the International Practical Kelvin Scale are obtained by adding 273.15 to values on the International Practical Temperature Scale.

[3] A complete list of the calibration services offered is contained in the NBS Test Fee Schedule as published in the Federal Register. Copies of separates covering temperature measuring instruments are available from the NBS.

Liquid-in-glass thermometers of the etched stem type are calibrated from -190° up to as high as 500 °C. The limitations of these thermometers are such that in the range 0 to 100 °C a limit of error of 0.01 deg can be achieved, but both above and below this range limits of error will increase considerably because of the relative instability of glass at high temperature and unsatisfactory physical properties of filling liquids at low temperature. An abrupt loss of accuracy, as shown in figure 1 occurs at -56 °C, which is the lower limit of use for mercury-thallium filled thermometers. Organic liquid fillings, which are used below this temperature, restrict the accuracy capabilities of low temperature thermometers of this type.

Optical pyrometers are calibrated in the range 800 to 4200 °C by comparison with the NBS standard optical pyrometer. Uncertainties inherent in the instruments themselves, as well as in the calibration process, increase the limits of error which can be assigned to calibrated optical pyrometers to 3 or 4 times the error in realizing the IPTS with the NBS standard. Limits of error usually assigned to certified instruments are 4 deg at 800 °C, 3 deg at 1063 °C, and gradually increasing to 40 deg at 4200 °C. In order to realize these accuracies with a calibrated pyrometer in practice, however, the body whose temperature is being measured must be radiating under blackbody conditions or its emissivity must be accurately known at the effective wavelength of the pyrometer. Tungsten strip lamps for use in checking or calibrating optical pyrometers are calibrated over the range from 800 to 2300 °C. Limits of error assigned in this calibration are 5 deg at 800 °C, 3 deg at 1100 °C, and 7 deg at 2300 °C.

4. Accuracy Improvement and Extension of Services

A continuing program of research directed toward the improvement and extension of the calibration services has always been in progress at the NBS in an attempt to keep abreast of the National needs. In recent years, however, rapid advances in science and technology are accelerating the demand for more accurate measurements over wider ranges. Accordingly, these research efforts have been expanded as rapidly as practical to not only improve existing standards and extend services into important areas not now adequately covered, but also, in so far as possible, to anticipate coming needs. Currently, activities are in progress throughout most of the temperature range from about 1 °K up into the 10,000 to 20,000 °K range of arc plasmas. Temperature measurements by accoustical interferometry [7], now under study, are expected to provide an accurate temperature scale bridging the gap between the lower limit of the NBS 1955 Scale and temperatures defined by the vapor pressure of liquid helium-4 [8]. Based upon this scale, services for the calibration of resistance type thermal elements over the temperature range 1° to 20 °K will be offered in the near future. At higher temperatures, furnaces are being developed which will extend the NBS thermocouple calibration capabilities from 1100° to about 2200 °C [9]. In addition, several other projects in progress are directed toward improving and extending the temperature scales used at the NBS as the basis for calibration services. Studies of the resistance-temperature relation of very pure platinum in the range covered by the NBS 1955 Scale, together with similar work in other national laboratories, are expected to provide a basis for international agreement to extend the IPTS to temperatures below its present lower limit at the oxygen point. The development of a high temperature resistance thermometer [10], which can replace the standard platinum thermocouple as the instrument for interpolation between fixed points on the IPTS at temperatures between 630.5 and 1063 °C, will greatly improve this part of the scale when international adoption is accomplished. Work on a photoelectric pyrometer [11], which has been in progress for several years, has reached a stage where this instrument will soon be used at the NBS to realize temperatures on the IPTS above the gold point. It is anticipated that accuracies attained in realizing this part of the IPTS will be improved by a factor of two or three through the use of the new instrument. At very high temperatures, spectroscopic techniques are being developed for the measurement of arc plasma temperatures of 10,000 to 20,000 °K [12].

The fruition of the work mentioned above will represent many man years of effort since unusual attention to detail is necessary at each stage of the progress of standards development. Much has already been accomplished, however, with much more to come in the next year or two.

5. References

[1] H. F. Stimson, International practical temperature scale of 1948, Text revision of 1960, J. Research NBS **65A** (Phys. and Chem.), No. 3, 139-145 (May-June 1961); also NBS Mono. 37 (Sept. 1961).

[2] H. F. Stimson, Precision resistance thermometry and fixed points, Book, Temperature, Its Measurement and Control in Science and Industry, **2**, 141 (Reinhold Publ. Corp., New York, N.Y. 1955).

[3] Calibration of platinum resistance thermometers, NBS Tech. News Bull. **45**, 62 (1961).

[4] W. F. Roeser and S. T. Lonberger, Methods of testing thermocouples and thermocouple materials, NBS Circ. 590 (Feb. 1958).

[5] J. F. Swindells, Calibration of liquid-in-glass thermometers, NBS Circ. 600 (Jan. 1959).

[6] H. J. Kostkowski and R. D. Lee, Theory and methods of optical pyrometry, NBS Mono. 41 (Mar. (1962).

[7] G. Cataland, M. Edlow, and H. H. Plumb, the determination of absolute temperatures from sound velocity measurements, Book, Temperature, Its Measurement and Control in

Science and Industry, 3, Pt. 1, 129 (Reinhold Publ. Corp., New York, 1962).

[8] H. van Dijk, The 1958 scale of temperatures for the liquid helium-4 region, Progress in Cryogenics, p. 123 (Haywood & Co., Ltd., London, England, 1960).

[9] D. B. Thomas, A furnace for thermocouple calibrations to 2200 °C, J. Research NBS, 66C (Eng. and Instr.), No. 3, 255-260 (July-Sept. 1962).

[10] J. P. Evans and G. W. Burns, A study of the stability of high temperature platinum resistance thermometers, Book, Temperature, Its Measurement and Control in Science and Industry, 3, Pt. 1, 313 (Reinhold Publ. Corp., New York, N.Y., 1962).

[11] R. D. Lee, The National Bureau of Standards photoelectric pyrometer, Book, Temperature, Its Measurement and Control in Science and Industry, 3, Pt. 1, 507 (Reinhold Publ. Corp., New York, N.Y., 1962).

[12] J. B. Shumaker and W. L. Wiese, Measurement of electron density and temperature in dense plasmas by application of line broadening theory, Book, Temperature, Its Measurement and Control in Science and Industry, 3, Pt. 1, 575 (Reinhold Publ. Corp., New York, N.Y., 1962).

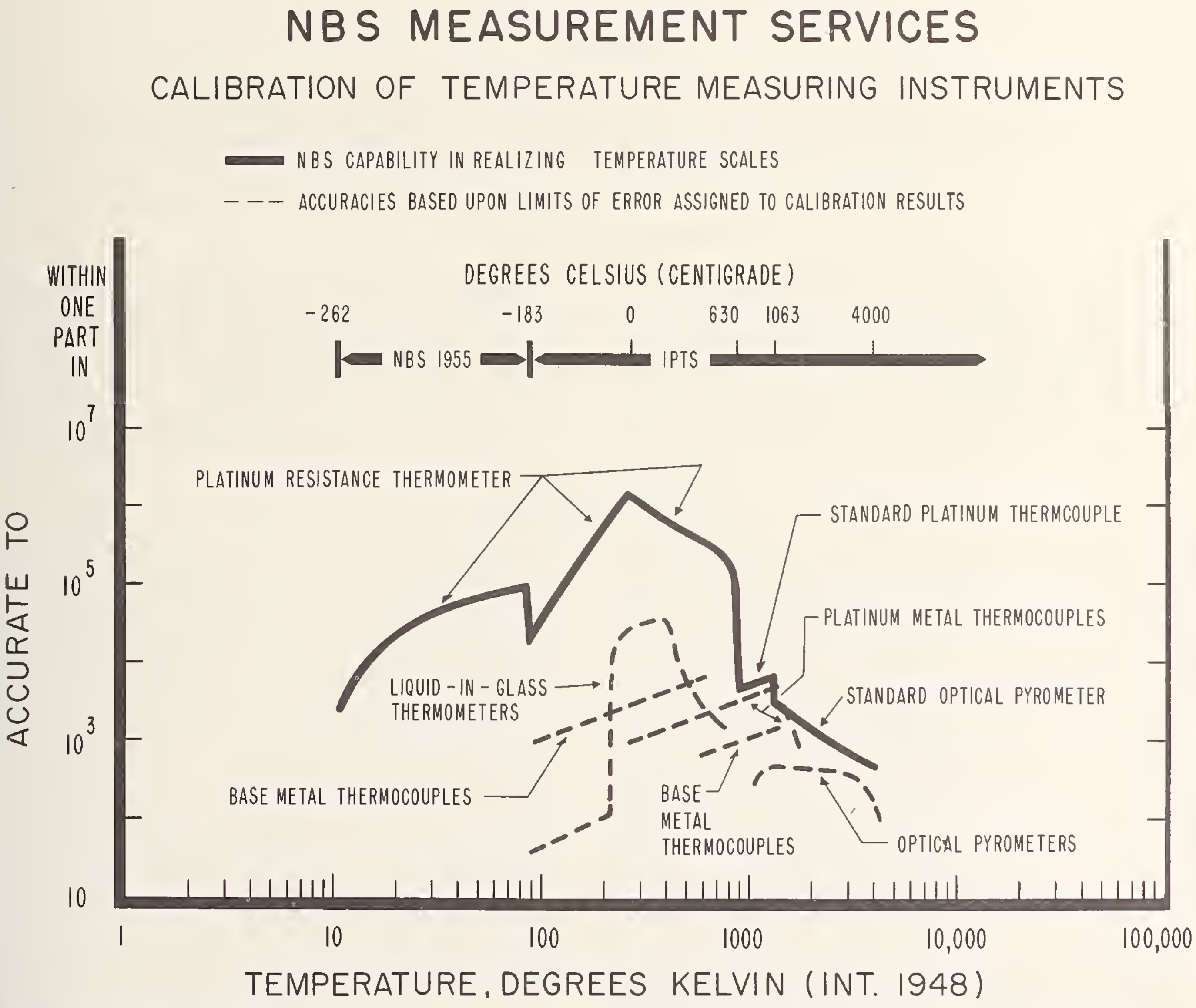


Figure 1. Accuracies attained at the National Bureau of Standards in reproducing the International Practical Temperature Scale and the NBS Provisional Scale of 1955 are represented by the solid lines. The broken lines show accuracies attainable with calibrated thermometers of selected types.

NBS Miscellaneous Publication 248
Proceedings of the 1962 Standards Laboratory Conference

Session 1. National Bureau of Standards Service to Industry

Paper 1.5. Low-Frequency Electrical Calibrations at the National Bureau of Standards

F. L. Hermach*

Charts are presented to show the range and accuracy of the National Bureau of Standards calibrations of standards of resistance, reactance, voltage, current, power and energy, from direct current through 30 kilocycles per second. The chains of measurements by which all of these calibrations are related to the basic NBS standards of voltage and resistance are also shown.

1. Introduction

There are many electric and magnetic quantities to be measured, such as resistance, voltage, and flux density. Measurements of most of them are made over many decades of both magnitude and frequency, with a really amazing variety of standards and instruments. It is certainly impractical to offer a calibration service at NBS to cover all of these variables. NBS concentrates instead on the calibration [1][1] of a few types of standards of the highest stability and accuracy. Because of this stability, the scientist and engineer can then verify the accuracy of his measurements with only infrequent periodic calibrations of these standards.

2. NBS Standards

Figure 1 shows the major electrical standards which are used by NBS in the calibration program. The lines indicate the major relationships between them. For clarity, some minor relationships and the kinds of calibrations performed are not shown.

The internationally accepted prototype standards of mass, length, and time are given in the top row of the figure. The meter is now defined as a certain number of wavelengths of the orange-red line of krypton 86 and the meter bar is a working, rather than a prototype, standard. [2][2] The second is now defined as a certain fraction of the tropical year 1900, and stable oscillators serve as working standards of its reciprocal (c/s). The kilogram is still the mass of the prototype Pt-Ir standard.

Two experiments are performed at NBS to determine the basic electrical units in terms of these three standards and two measured constants, the speed of light in vacuo (c) and the acceleration of gravity (g). They are simple in principle but extremely difficult and involved in practice, because of the accuracy required. One experiment consists in constructing a capacitor from gage bars and computing its capacitance in electromagnetic units (about 1 pf) from the length of the bars and the speed of light. With suitable bridges the step up is made to two 10,000-pf capacitors, across to two 10,000-ohm resistors at 1592 c/s, and down to nominal 1-ohm resistors. [3] This is done because 1-ohm Thomas-type resistors are the most stable impedance standards known. The average resistance of a group of such resistors serves to maintain the ohm at NBS between such absolute determinations.

The other experiment consists in "weighing the ampere" with our current balance, and is thus based directly on the definition of the ampere in terms of the force between current-carrying conductors. [4] One conductor (coil) is suspended from one arm of a balance and the force is compared with force of gravity on a known mass. The voltage drop in a 1-ohm standard resistor carrying this

---

*Electrical Instruments Section, Electricity Division, National Bureau of Standards, Washington, D. C.

[1] "Calibration" is used here as the process of making appropriate measurements to determine the correct value of a standard.

[2] Figures in brackets indicate the literature references at the end of this paper.

measured direct current is then used to determine the emf of a group of saturated standard cells, which, in turn, maintain the volt at NBS.

Direct-reading ratio sets were originated at NBS to calibrate resistors by the substitution method, and to step up and down on the resistance scale. [5] No line leads to them in figure 1 because their accuracy depends on ratios of resistors, not on the unit of resistance. Standard cells are calibrated by connecting the known and unknown cells in opposition and measuring the small voltage difference with a low-range thermofree potentiometer. Potentiometers and volt boxes of the highest quality are calibrated with universal and direct-reading ratio sets. With these the user can then extend the d-c voltage scale very accurately. Alternating current-direct current transfer instruments are comparators for determining the equality of a-c and d-c voltages, a-c and d-c currents, and a-c and d-c powers.[6] At NBS they serve chiefly to determine how well other ac-dc comparators do this. Since such comparators are very stable, the user can make accurate a-c measurements with his known d-c standards and ac-dc transfer techniques. Alternating current bridges extend the scale of capacitance and inductance measurements, but NBS calibrates only the more stable capacitors and inductors, with which bridges can be checked. Alternating current voltage dividers and potentiometers are used to determine the ratios (expressed as complex numbers) of the voltage and currents of the NBS standard voltage and current transformers. In principle they are also independent of units, but in practice some of the component impedors are measured with the d-c ratio sets and a-c bridges.

The standard watthour meters and standards for magnetic measurements depend on the other standards as shown in figure 1.[7] [8]

3. Accuracy Charts

Figures 2 through 6 show the accuracies attainable at NBS in the calibration of an "ideal" standard; that is, one which is perfectly stable and definite, and is free from influences of temperature, humidity, etc. In each figure, the accuracy or uncertainty is considered as the "limit of error" which would be exceeded on only rare occasions. It includes the calculated imprecision and the estimated residual systematic errors of the NBS working standards and the calibration process, but not systematic errors in the unit established and maintained at NBS.

The certified accuracy of an actual standard is generally considerably poorer than the values shown on the charts because of three significant factors. *First* an actual standard is not perfectly definite, and is subject to external influences, all of which degrade the accuracy of the calibration. In the relatively few measurements that can be taken on any one standard, these influences cannot be evaluated precisely. *Second* with few exceptions, such as Thomas-type resistors and saturated cells, the values are certified in the theoretical or absolute units. It is very unlikely that the uncertainty in the unit maintained at NBS will exceed the value shown on each chart. *Third* since long time stability is such an important factor, an allowance is made in many cases for expected changes in the standard for one year, based on NBS experience with the particular standard or others of the same type. In other cases, the uncertainty of the calibration itself is stated, often with a separate estimate of the expected stability.

NBS intends to maintain working standards of the same quality and accuracy at both the Boulder, Colo. and Washington, D. C. laboratories, so that, with few exceptions, normal calibration services at low frequencies can be obtained at either location.

Figure 2, for resistors, shows the familiar "accuracy triangle," with the peak at the value of the 1-ohm basic standard. Note that only decimal multiples and submultiples of 1-ohm are shown; NBS does not usually offer calibrations of other values nor does it maintain an accurate resistance measuring facility as such. The resistances of Thomas-type 1-ohm resistors can be certified to 2 ppm in terms of the ohm maintained at NBS, including the first and third factors; others are normally certified to 20 ppm, including all three factors mentioned. Direct-reading and universal ratio sets, not shown in figure 2, can be calibrated to a few ppm.

It is interesting to note that for a short while during the absolute measurements, the 10-, 100-, 1000-, and 10,000-ohm NBS standard resistors are necessarily known to better absolute accuracy than are the 1-ohm resistors used to maintain the unit, because of small errors inherent in the step-down procedure. However, this is only temporarily true because of the much better long-time stability of the 1-ohm resistors.

The chart for capacitance measurements, figure 3, should be three-dimensional, with frequency as another independent variable. One would expect, in a rough way, an "accuracy cone" with the peak at 1 pf and 1592 c/s, where the best measurements of the computable capacitor have been made. Figure 3 shows a few of the contours. The research bridge is not available for calibrations, but some of the standards in the other bridges are calibrated with it. Capacitors of suitable quality are normally certified to 0.03 percent depending on range and frequency. This includes all three of the accuracy factors given.

The large air-cored inductors formerly used at NBS for determining the ohm cannot now be used with suitable accuracy because of nearby magnetic materials and interfering fields. Less stable secondary standard inductors are calibrated with a Maxwell-Wien bridge and standard capacitors, and are used as working standards with an inductance bridge (fig. 4). Nevertheless, high-grade inductors can be certified to 0.03 percent including all three factors.

The chart of voltage measurements is complicated not only because of applied frequency but also because NBS does not offer a voltage calibration service as such. Only the d-c voltage of a standard cell is available directly (and is certified to 0.6 ppm in terms of the volt maintained at NBS). The scale is extended in voltage and frequency by the calibrations of ratio standards and ac-dc transfer standards. With these and a standard cell the user can then make accurate d-c and a-c measurements. The ratio standards (potentiometers, fixed resistance dividers, and transformers) are generally used with respect to some reference voltage, such as the emf of a cell or a nominal 120-v line, and have upper and lower voltage limits. The chart therefore shows the accuracy of the ratio V/V_r, where V is given along the abscissa. Note that this is the accuracy of the indicated ratio, not of a fraction of the applied voltage. (A 1000/1 volt box certified to 0.01 percent has one-tenth the uncertainty of a 1 ppm-of-input voltage divider at this same ratio). Potentiometers are normally certified to 0.01 percent or 0.005 percent, volt boxes to 0.01 percent of ratio, inductive voltage dividers to 0.001 percent of input-voltage at 1kc/s, and voltage transformers to 0.05 percent of ratio at 60 c/s. Alternating current-direct current transfer instruments are certified to 0.01 percent from 20 c/s to 20 kc/s. Since these are ratio devices and comparators, the second factor (the uncertainty of the NBS unit) does not enter directly.

The chart for current measurements, figure 6, is similar to figure 5. Here NBS offers only resistance, ratio-calibration and ac-dc transfer services. A chart for power measurements would include current and voltage standards and the NBS standard transfer wattmeter, which is known to 0.005 percent at 60 c/s and 0.1 percent at 2 kc/s.

4. Summary

These charts show the accuracies presently attainable under the restictions given. In every case the available or certified accuracy is necessarily poorer because of the factors stated. NBS recognizes that in many areas improved accuracy is vitally needed and is striving to meet the demands by developing better standards and improving the accuracy of existing standards. Thus these charts should not be considered as any more permanent than the calibrations of the standards they depict.

The author thanks Ralph Kotter and Chester Peterson for figures 2, 3, and 4, and for the ideas they contributed in discussions of this subject.

5. References

[1] F. B. Silsbee, Systems of electrical units, J. Research NBS **66C** (Eng. and Instr.), No. 2, 137-178 (Apr.-June 1962); also NBS Mono. 56 (Sept. 1962).

[2] Actions of 11th Conference on Weights and Measures, NBS Tech. News Bull. **44**, 199 (Dec. 1960).

[3] R. D. Cutkosky, Evaluation of the NBS unit of resistance based on a computable capacitor, J. Research NBS **65A** (Phys. and Chem.) No. 3, 147-158 (May-June 1961).

[4] R. L. Driscoll and R. D. Cutkosky, Measurement of current with the NBS balance, J. Research NBS **60**, 297 (1958) RP2846.

[5] J. L. Thomas, Precision resistors and their measurements, NBS Circ. 470 (Oct. 1948).

]6] F. L. Hermach, Thermal converters as ac-dc transfer standards, J. Research NBS **48**, 121 (1952) RP2296.

[7] A. W. Spinks and T. L. Zapf, Precise comparison method of testing ac watthour meters, J Research NBS **53**, 95 (1954) RP2521.

[8] R. L. Sanford and I. L. Cooter, Basic magnetic quantities and the measurement of magnetic properties of materials, NBS Mono. 47 (May 1962).

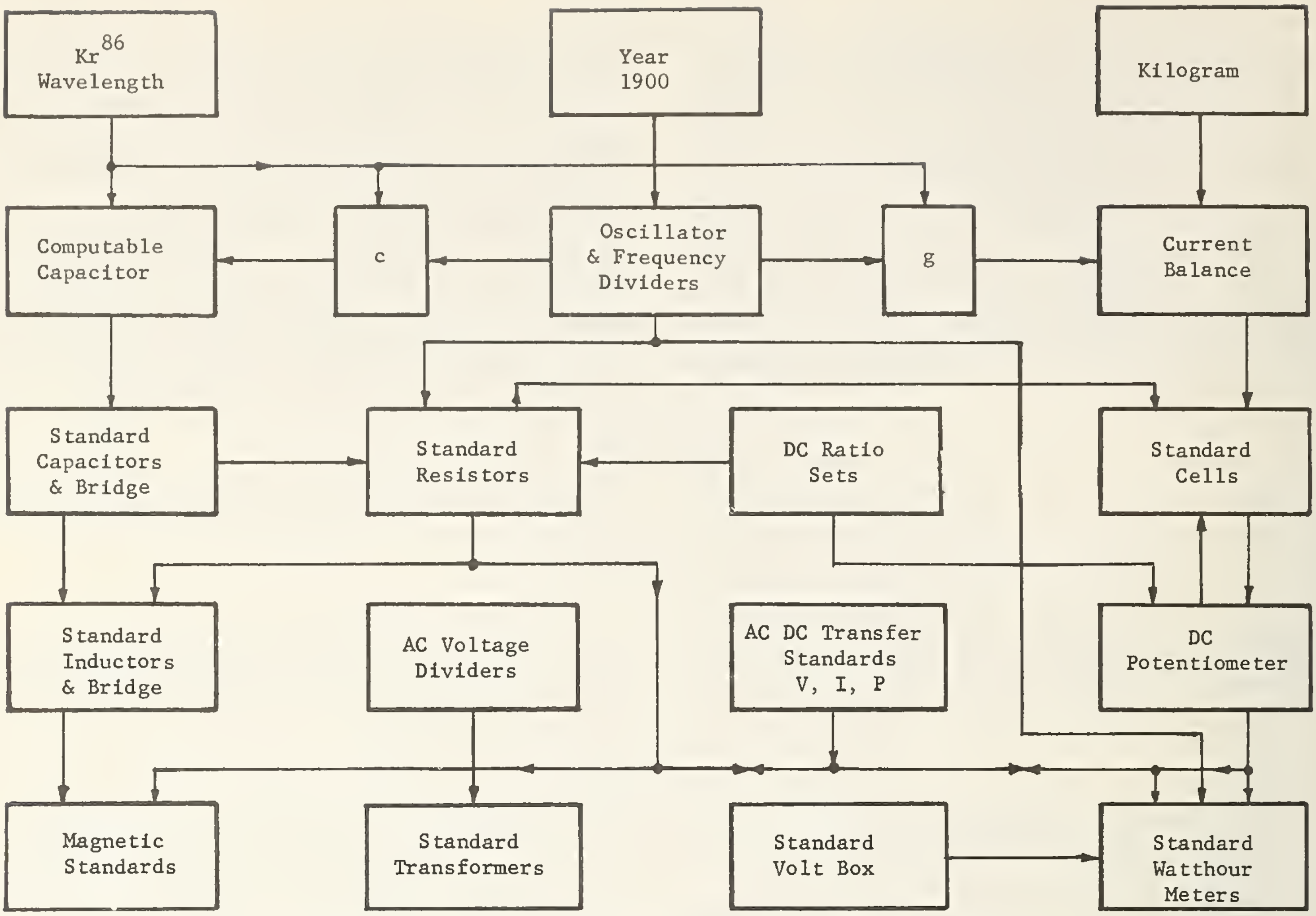


Figure 1. Electrical standards used in calibration program at NBS.

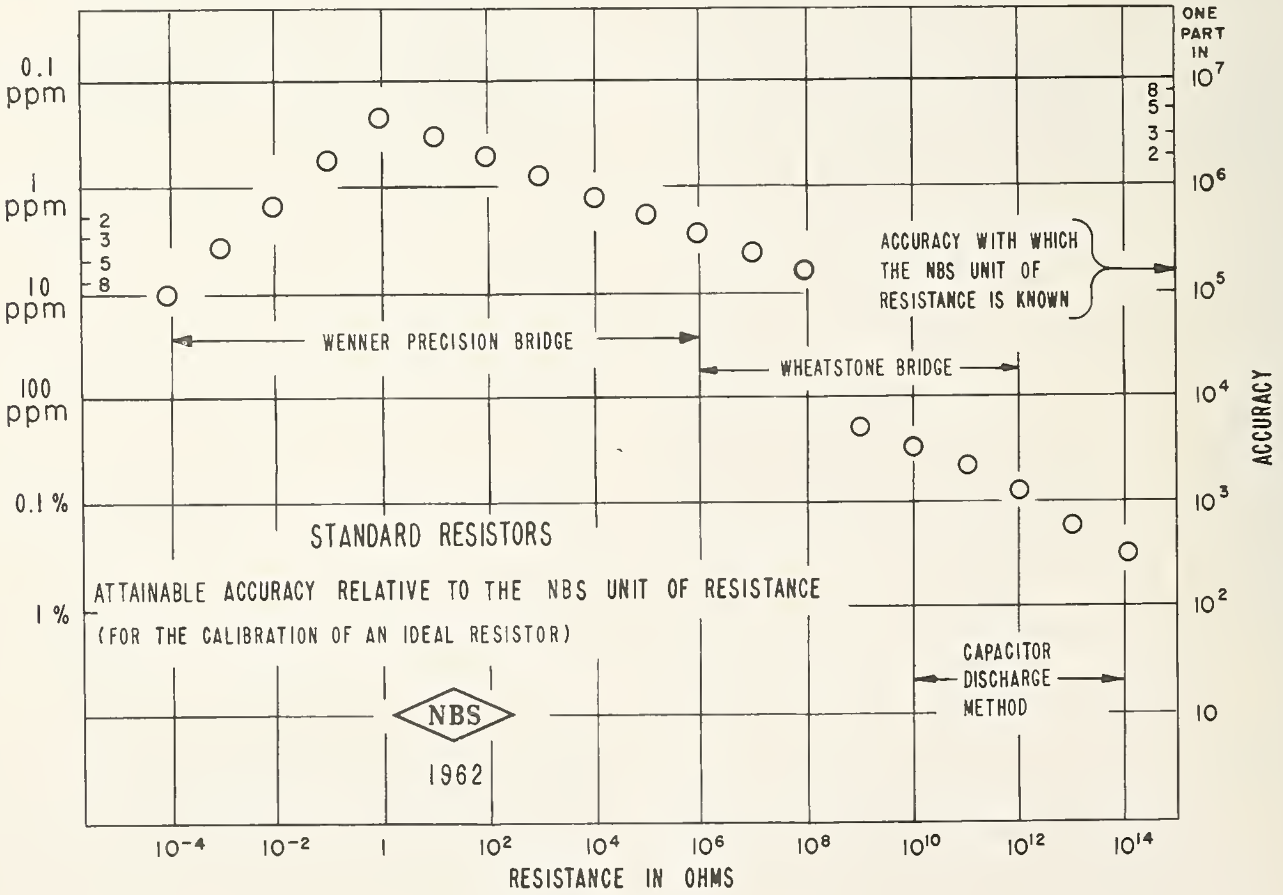


Figure 2.

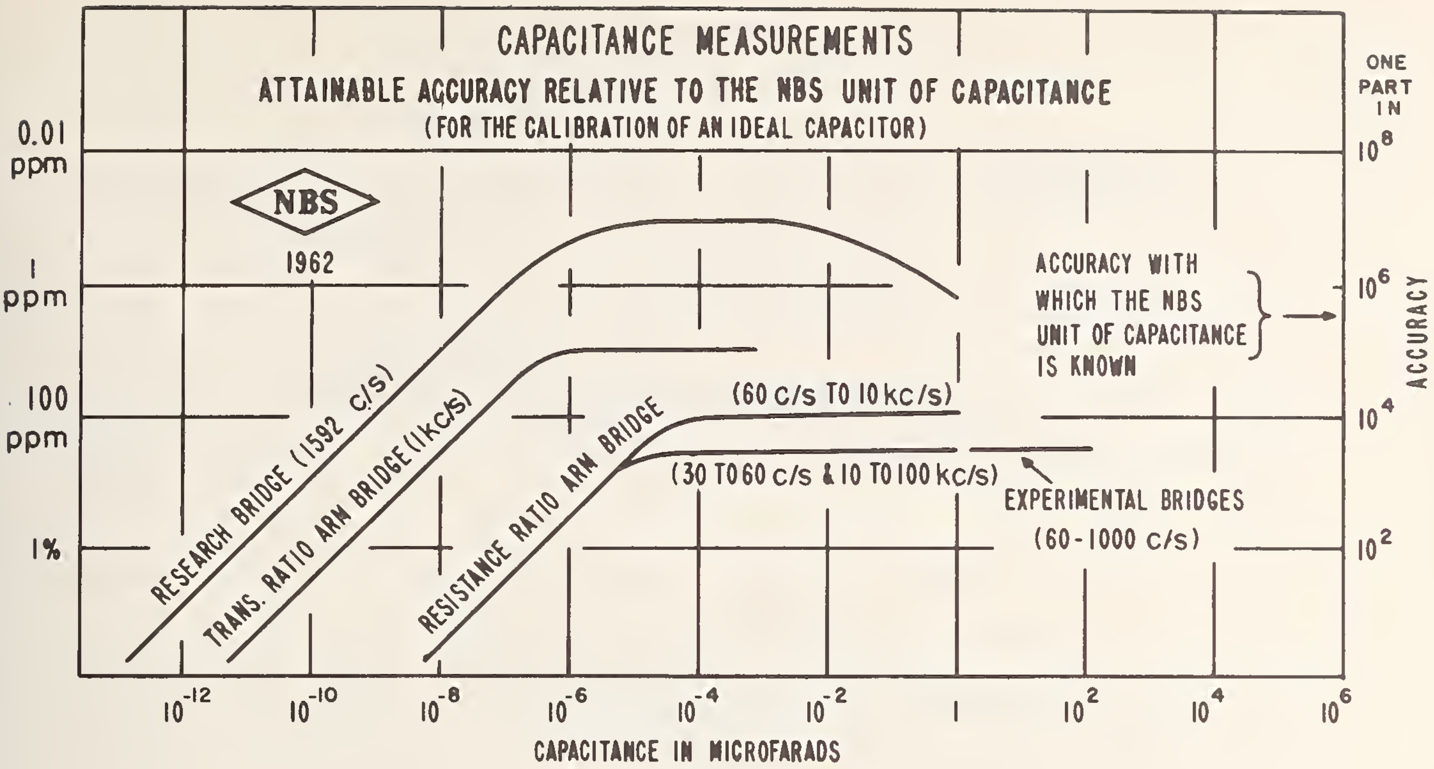


Figure 3.

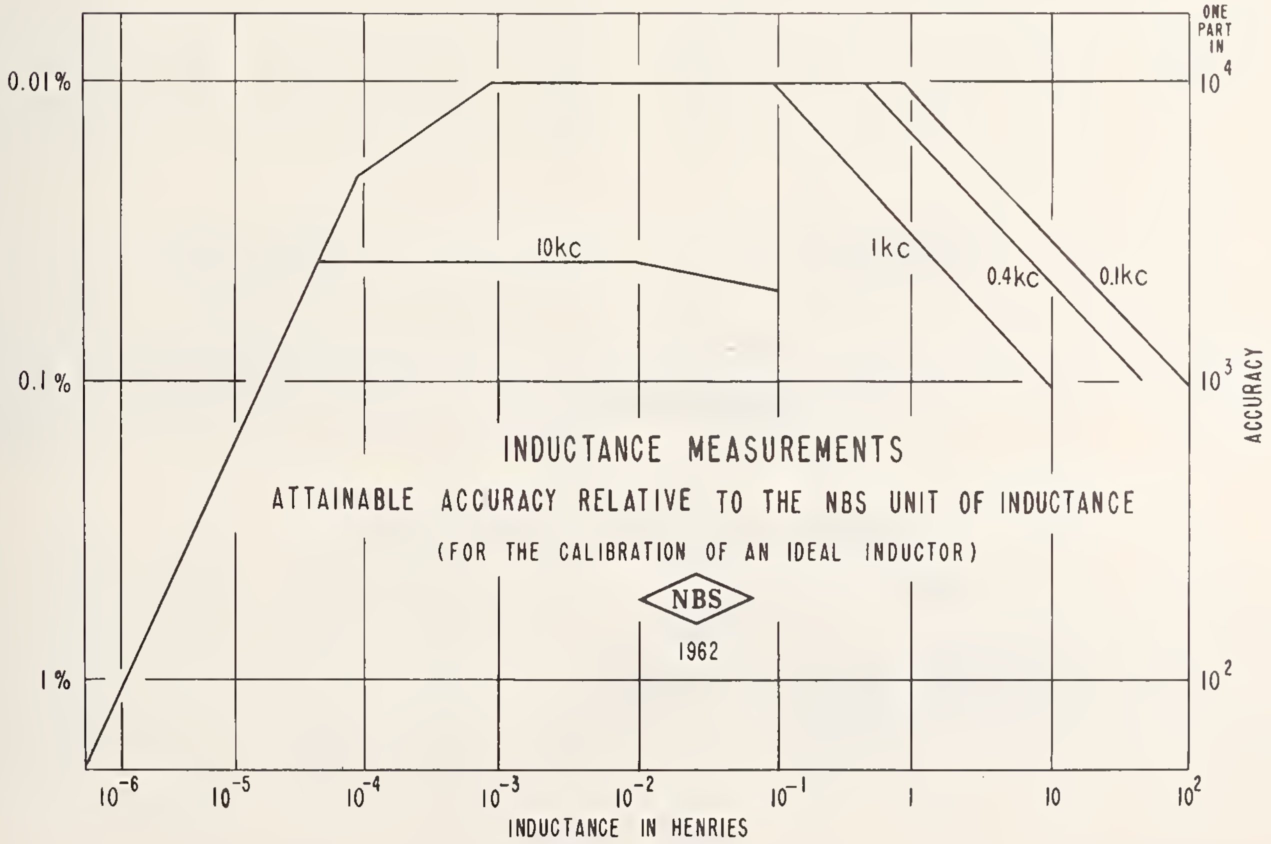


Figure 4.

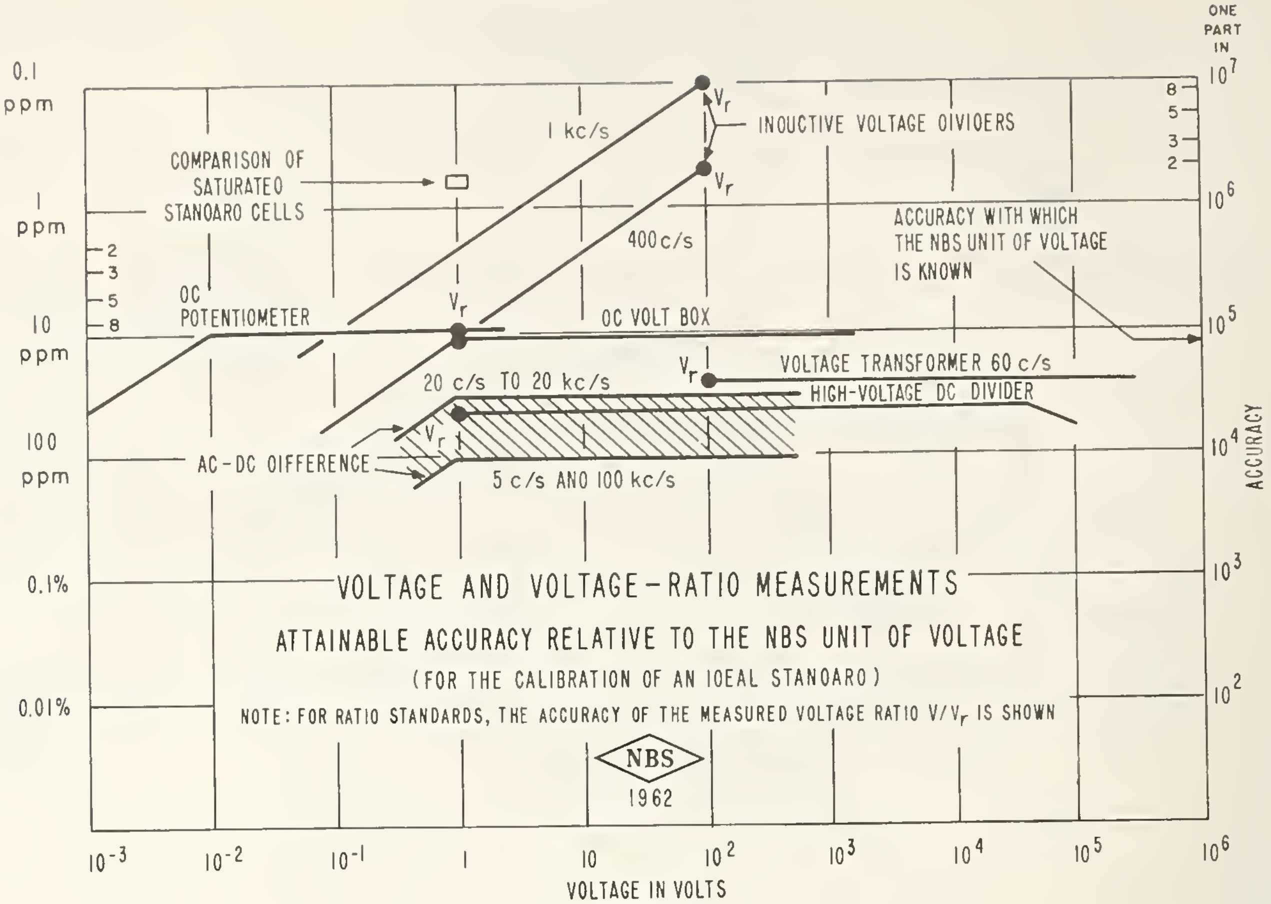


Figure 5.

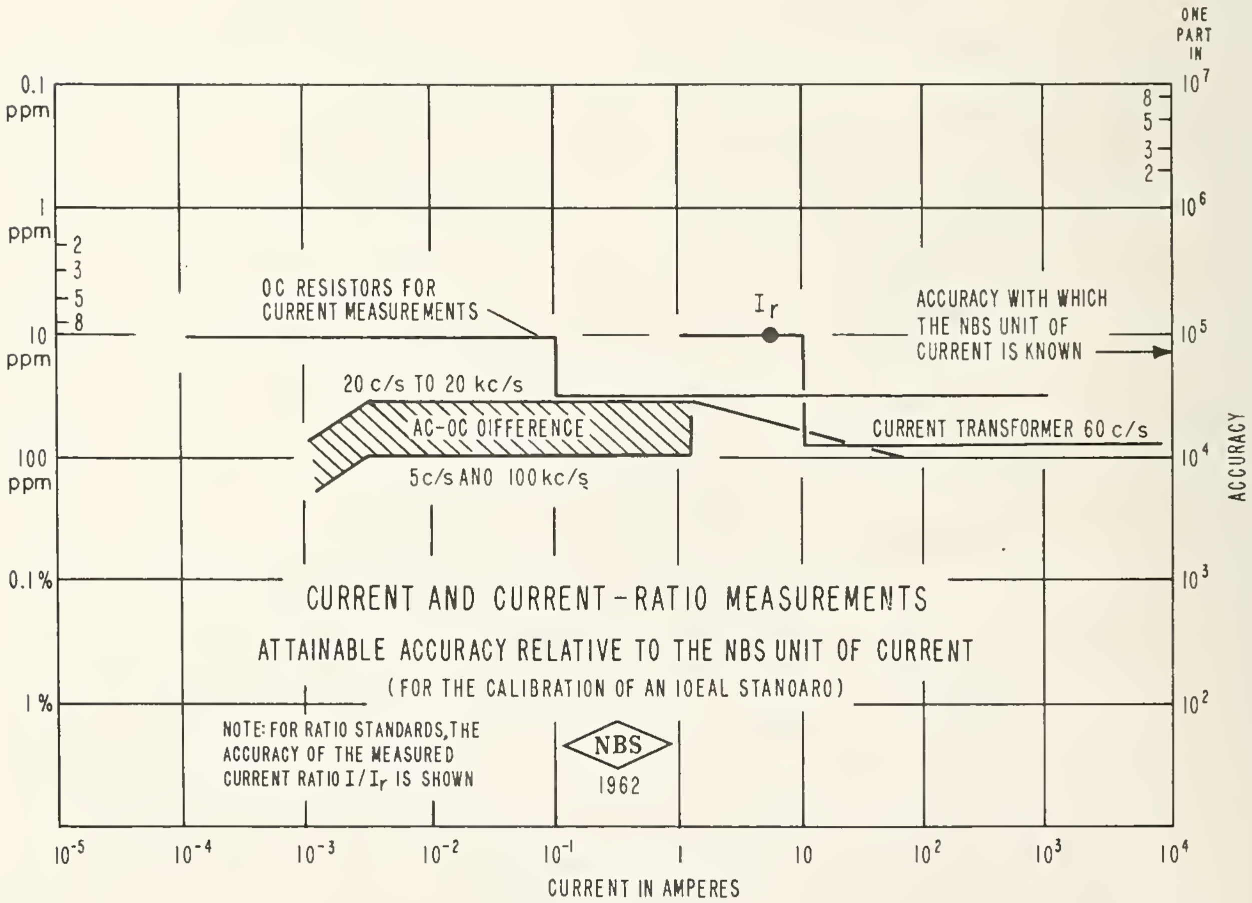


Figure 6.

Session 1. National Bureau of Standards Service to Industry

Paper 1. 6. Frequency and Time Calibration Services at the Boulder Laboratories of the National Bureau of Standards

A. H. Morgan*

The present frequency and time calibration facilities and methods and techniques are presented, followed by some of the future plans in this area. Also included is a description of the present NBS VLF, LF, and HF standard broadcasts and some discussion of the results obtained. Some data presented are on various measurements made on highly stable quartz oscillators.

1. Introduction

Until rather recently the standard high frequency (HF) and time signal broadcasts of WWV and WWVH [NBS, 1960] have provided frequency and time calibration facilities, for the United States (and much of the rest of the world) of sufficient accuracy to meet practically all of the needs in this area. However, the recent rapid advances in many fields of science have made demands for increased accuracies that are not possible to achieve with the HF broadcasts. This situation was foreseen as early as 1956 when the standard radio broadcasts were begun by the NBS in the low frequency (LF) band at 60 kc/s (using the call letters KK2XEI, later changed to WWVB) and located near Boulder, Colo.

In response to many requirements involving means for synchronizing widely separated clocks [Morgan, 1959] and providing accurate time signals worldwide, [Watt, Plush, Brown, and Morgan, 1961] a study was undertaken at NBS which indicated that an optimum frequency [Watt and Plush, 1959] for this purpose was in the VLF band at about 20 kc/s. Accordingly, in April 1960, standard radio station WWVL, near Sunset, Colo., was put into operation at 20 kc/s. [Shoaf, 1962.]

The signals of stations WWVL and WWVB have been received in nearly all parts of the Continental United States and in Canada, and the 20 kc/s signal has been received as far away as New Zealand, [Crombie, 1960] using a phase-lock receiving system. The carrier frequencies of these stations are controlled by the U.S. Working Frequency Standard (USWFS) and are stable to about ± 3 parts in 10^{11}; WWVB is directly controlled while WWVL is phase-locked [Fey, Milton, and Morgan, 1962] by means of a 50 Mc/s radio link, which was put into operation in October 1961. This is a first in this field.

In addition to the standard radio broadcasts, there are some special measurements that may be made at the Boulder Laboratories that require facilities not available to many Standards Laboratories. This includes measurements of highly stable frequency sources, such as quartz oscillators, rubidium (Rb) gas cells, cesium (Cs) beam standards, etc., and measurements of the power spectra of frequency sources.

2. Frequency and Time Standards

The present standard of time [Comité International des Poids et Mesures, 1957] is the Ephemeris Second and it is defined as the fraction $\frac{1}{31,556,925.9747}$ of the tropical year for January 0, 1900 at 1200 hours, Ephemeris Time (ET). It is determined by astronomical means to an uncertainty of a few parts in 10^9 in a period of a few years. Because time and frequency are inversely related by definition, standards of frequency and time are fundamentally the same. To make the standard quickly available, a relationship between the atomic transition frequency of cesium (Cs) and ET was

---

*Chief, Frequency and Time Dissemination Section, Radio Physics Division, National Bureau of Standards, Boulder Laboratories, Boulder, Colo.

found to be desirable. This was done [Markowitz, Hall, Essen, and Parry, 1958] in 1958 and the value given, which is the only one available, is that there are 9,192,631,770 ± 20 cycles of Cs per Ephemeris Second. The uncertainty of about two parts in 10^9 in this value should, in the strictest sense, be transferred to the frequency derived from it. However, for all but perhaps the "purists," this is generally not a necessity and, for most high precision frequency measurements, not desirable. This is because it then is possible to take full advantages of the remarkable stability and accuracy of the Cs standards, [Mockler, Beehler, and Snider, 1960] where accuracy refers to the closeness with which the ideal transition frequency of Cs is realized, which is one to three parts in 10^{11}. It is quite generally agreed now that astronomical time will never match atomic time in regard to stability, accuracy, and availability.

The Consultative Committee [Comité Consultatif pour la Definition de la Seconde, 1961] for the Definition of the Second of the International Committee on Weights and Measures (ICWM) met in Paris in April 1961 to consider the question of defining time in terms of an atomic standard. It seems quite probable that this may be done when the ICWM meets again in 1966.

3. Frequency and Time Standards at NBS

3.1. Uniqueness of Frequency and Time Standards

There are at least three ways in which frequency and time standards are unique. They: (1) are not permanent (such as a kilogram cylinder which constitutes the standard of mass) but must be continuously reconstructed and checked: (2) can be measured with higher precision than any other physical or electrical quantity, and (3) can be made widely available by means of radio signals.

3.2. U.S. Frequency Standard

The NBS has developed and maintains for the United States [NBS, 1960] and its outlying bases, accurate and precise cesium beam frequency standards [Mockler, Beehler, and Snider, 1960] that constitute the U.S. Frequency Standard (USFS). They are among the most accurate, stable and reproducible atomic frequency standards in the world.

In addition, an ammonia beam Maser [Barnes, Allan, and Wainwright, 1962] has been recently developed at NBS. It has three properties that make it quite unique as a spectrum analyzer, [Barnes and Heim, 1961]: (1) it has a natural line width of the transition frequency that indicates a Q factor of about 10^7, which is about equivalent to that of the best quartz crystals at room temperature or above: (2) the oscillation frequency is at K band without frequency multiplication, and (3) it is inherently a very low noise device.

3.3. U.S. Working Frequency Standards

In addition to these high quality standards, the U.S. Working Frequency Standards (USWFS) are maintained and periodically calibrated in terms of the USFS. The former, consisting of several highly stable and precise commercial quartz and atomic standards, generate output frequencies which are in general constant to about three parts in 10^{11}, and are used to: (1) control or steer all of the NBS standard radio broadcasts (WWV, WWVH, WWVB, and WWVL): (2) distribute standard signals throughout the Boulder Laboratory: (3) control the microwave (MW) frequency standard in the Electronic Calibration Center: (4) calibrate equipment sent to Boulder Laboratory, and (5) control precise clocks.

4. Standard Radio Broadcasts

4.1. High Frequency Broadcasts of WWV and WWVH

The NBS maintains standard frequency and time signal broadcasts [NBS Miscellaneous Publication 236, 1960] at HF using stations WWV, Beltsville, Maryland, and WWVH, Maui, Hawaii. Both stations provide the following technical services: (1) standard radio frequencies: (2) standard audio frequencies: (3) standard time intervals: (4) standard musical pitch, and (5) time signals. All of these signals at each station are derived from a common master oscillator whose frequency is based on the USFS and whose daily stability is better than ± 5 parts in 10^{11} at WWV and one part in 10^{10} at WWVH.

4.1.1. Frequency Offset From USFS

Beginning in January 1960 the master oscillators at WWV and WWVH have been intentionally offset from the USFS by a small, but precisely known, amount in order to reduce the departure of the time signals broadcast from UT2 time. In 1962 the offset was -130 parts in 10^{10} and will be the same in 1963. From January 1960 to January 1962 it was -150 parts in 10^{10}. It is expected that the offset may be left unchanged throughout the calendar year even though UT2 time is subject to unpredictable changes that are easily detected at the level of precisions involved.

4.1.2. Corrections to Carrier Frequencies Broadcast

For these reasons, corrections to the carrier frequencies as broadcast are determined by the NBS with respect to the USFS by means of the NBS standard low frequency (LF) and very low frequency (VLF) radio broadcasts, which are described below. These values are published monthly [WWV Standard Frequency Transmissions, 1958] in the Proc. IRE and are available in back issues to May 1958 with the data extending back to December 1, 1957. The data are given to one part in 10^{11} with an uncertainty of five parts in 10^{11}.

4.1.3. Standard Time Intervals and Time Signals

Highly precise time interval and time marking signals, consisting of five cycles of 1,000 c/s at WWV and six cycles of 1,200 c/s at WWVH, are broadcast as a pulse at intervals of precisely one second. At the 59 sec. of each minute the pulse is omitted but two pulses separated by 0.1 sec. are emitted on the 60 sec. To precisely mark the beginning of each 5 min. period beginning on the hour, an audiofrequency modulation of 600 c/s is keyed on at both stations and continues for 3 min. at WWVH and 2 min. at WWV. The rest of the first 5 min. period at WWVH is silent except for the time pulses and the International Morse code time announcements occurring during the first half of the last minute. At WWV, the third minute has a time code modulation, the fourth and fifth have only the time pulses except for the time announcements made during the last half of the fifth minute. The next 5 min. period starts with a 440 c/s tone modulation but is identical otherwise with the first. These 5 min. periods alternate during each hour, except for the scheduled silent periods. At WWVH this is from 15 to 19 min. past each hour, and at WWV it is from 45 to 49 min. past each hour. Also, WWVH has a scheduled silent period from 1900 to 1930 UT, daily.

4.1.4. Step Adjustments of Time Interval and Time Signals

The time signals and time intervals are kept in close agreement with UT2 time [Markowitz, unpublished] by making step adjustments, of precisely known amounts when necessary. Since 1959 the step adjustments were made as follows; at 0000 UT on date given:

(1) December 16, 1959, retarded by 20 milliseconds (ms)
(2) January 1, 1961, retarded by 5 milliseconds (ms)
(3) August 1, 1961, retarded by 50 milliseconds (ms)

Future adjustments will be made when necessary in steps of precisely 50 ms on the first day of the month following the one in which the transmitted time departs from UT2 time by more than 50 ms.

4.1.5. Coordination of Time Signals With Other Nations

See appendix 3 for details on this.

4.1.6. Time Code on WWV

See appendix 4 for details on this.

4.2. LF and VLF Standard Broadcasts of WWVB and WWVL

NBS provides standard frequency broadcasts in the LF band (WWVB at 60 kc/s) and in the VLF band (WWVL at 20 kc/s). The former (with call sign KK2XEI) was begun in June 1956 and the latter in April 1960. Both carrier frequencies are continuously maintained and are controlled by the USWFS; WWVB is directly controlled and WWVL by a very unique and the first of its kind, remote phase-control system [Fey, Milton, and Morgan, 1962]. This corrects for the frequency and phase changes due either to the controlling oscillator, the transmitter or the antenna system, so that the transmitted carrier frequency at 20 kc/s is essentially as stable as that of the USWFS (± 3 parts in 10^{11}).

4.2.1. Frequency Offset

The carrier frequencies of WWVB and WWVL, as transmitted, are also offset from the USFS by the same amount as those of WWV and WWVH so (1) that when precise time signals are added to them, they will be in step with those of the other nations, as described above, and, (2) to minimize the inconvenience to the users of UT2 time.

5. Measurements on Stable Frequency Sources

5.1. General

No known radio frequency source is truly monochromatic regardless of how stable it may appear when measured over a given time interval. In the strictest sense, then, a source does not have a single frequency but a spectrum of frequencies, the latter depending on the manner in which the source is modulated and the characteristics of the unwanted modulating signals. Unfortunately, there is at present no single quantitative way of completely describing [Strandberg, 1960] a frequency source and so several measurements are necessary to completely determine its performance characteristics. These include stability measurements taken over various periods of time from milliseconds (ms) to hours, and up to several days; also, power spectrum measurements may often be necessary or useful.

5.2. Frequency Stability Measurements

For intervals of time shorter than about one day, the stability of a quartz oscillator is determined by effects other than the aging of the crystals. This includes unwanted modulation, noise and temperature effects, component instabilities, effects of voltage variations, etc., whose individual or combined effects are not predictable to any useful degree. Thus, it is necessary to make stability measurements over the time intervals of interest in terms of a source whose stability is better than the unknown. Usually, at BL this is either a commercial rubidium (Rb) gas cell standard or one of the NBS ammonia Masers, the choice depending on the period over which the source is measured. For intervals shorter than a few seconds the Maser is used, as is also true in the spectrum measurements described below. For long term stability measurements, a Rb standard is used but it is calibrated periodically in terms of the USFS.

5.3. Power Spectrum Measurements

To make use of some of the unique properties of the ammonia Maser, a spectrum analyzer system [Barnes and Heim, 1961] was developed. The source to be measured is multiplied in frequency to a value near that of the Maser frequency (23.870 Gc/s), mixed with that of the Maser, and the resulting beat note is analyzed with a narrow-band (3 c/s) crystal filter to obtain the power spectrum. The latter shows the relative energy distribution of the various frequency components of the source and is thus a measure of the spectral purity of the signal.

The process of frequency multiplication, as is well known and used in the Armstrong FM system, has certain definite effects on the power spectrum of the signal. All stable frequency sources have very little, if any, residual AM due to their inherently high amplitude stability, but do contain FM signals of varying degree. The multiplication in frequency of such sources also multiplies the basic FM modulation index (β) by the factor of multiplication. This means the sidebands at the carrier frequency are increased in amplitude by the frequency multiplication process with the usual reduction in the amplitude of the carrier caused by the increased β, where β is defined as the ratio of the modulating frequency to the maximum frequency deviation of the carrier.

In this system, the FM modulation index is less than one and the large multiplication factor enhances the FM sidebands by about 73 db. This greatly facilitates the spectral measurements of the relatively "purer" sources. Further details are given in Barnes and Heim 1961.

6. Future Plans for Improving and Extending the Services

6.1. Improved Signals 60 kc/s and 20 kc/s

Construction is underway on a new LF and VLF station, at Ft. Collins, Colo., so as to provide early in 1963 improved services by these broadcasts. Station WWVB (60 kc/s) will radiate a carrier frequency of about 7 kw and will also have precise time signals consisting of five cycles of 1,000 c/s, transmitted once per second. They will be kept in agreement with the WWV/WWVH time signals, as transmitted, to within ± 1 ms, initially, with plans to increase this to less than 10μsec at a later date. These signals should be generally useful over much of the Continental United States and, by use of coherent pulse reception or similar techniques [under development at NBSBL), at much longer distances.

Station WWVL (20 kc/s) will provide a carrier frequency suitable for frequency calibrations, using phase-lock receiving techniques, over most of the world. The radiated power will be about 1 kw. It will also be an experimental facility for testing methods [Watt, Plush, Brown, and Morgan, 1961] of providing precise time signals, for worldwide clock synchronization from a single station.

6.2. Portable Clocks

In order to use the time signals from a radio station to accurately synchronize remote clocks to the master clock at the station, it is necessary to know the group delay times of the signals. For microsecond timing, the calculation of delay times over long radio paths is not of sufficient accuracy because of the uncertainty with which the parameters effecting the propagation velocity are known. This is true even for signals propagated by the ground wave over distances exceeding a few hundred miles. Therefore, it is necessary to measure the delay time over each radio path. The easiest and most accurate way to do this, and the initial accurate synchronization of clocks, [Morgan, 1959], is by means of highly stable and accurate portable clocks [methods under development at NBSBL].

It is planned to measure delay times of certain special radio paths and synchronize the clocks with those at stations WWVB and WWVL. Also, to provide very accurate synchronization of the latter stations with WWV and WWVH.

7. References

Barnes, J. A., and L. E. Heim, A high resolution ammonia-Maser spectrum analyzer, IRE Trans. Instr. **I-10,** 4-8 (June 1961).

Barnes, J. A., D. W. Allan, and A. E. Wainwright, The ammonia beam Maser as a standard of frequency, IRE Trans. Instr. **I-11,** No. 1, 26-30 (June 1962).

Comité International des Poids et Mesures, Proces-Berbaux des Seances, 1956 Session, Ser. 12, 25 (Gauthier-Villars, Paris, 1957).

Comité Consultatif pour la Definition de la Seconde, 2^e Session, Proc. (April 1961).

Crombie, D. D., Private communication (1960).

Fey, R L., J. B. Milton, and A.H. Morgan, Remote phase control of radio station WWVL, Nature **193,** No. 4820, 1063-1064 (Mar. 17, 1962).

Markowitz, Wm., R. G. Hall L. Essen, and J. V. L. Parry, Phys, Rev. Letters **1,** 105-106 (Aug. 1958).

Markowitz, Wm., Astronomical and atomic times, U.S. Naval observatory, Washington E. C. (unpublished).

Mockler, R. C., R. E. Beehler, and C. S. Snider, Atomic beam frequency standards, IRE Trans. Instr. **I-9,** 2, 120-131 (Sept. 1960).

Morgan, A. H., Precise time synchronization of widely separated clocks, NBS Tech. Note 22, (PB151381) (1959).

Morgan, A. H., J. A. Barnes, Short time stability of a quartz crystal oscillator as measured with an ammonia Maser, Proc. IRE, **47,** No. 10, 1782 (Oct. 1959).

Standard Frequency and Time Signals from WWV and WWVH, NBS Misc. Publ. 236 (revised, July 1, 1961).

NBS, National standards of time and frequency in the U.S., Proc. IRE **48,** No. 1, 105 (January 1960).

Shoaf, J. H., Standard frequency station WWVL (unpublished report) 1962.

Strandberg, M. W. P., Precise specifications of a frequency, The Microwave J. **3** No. 8, 45-50 (Aug. 1960).

U.S. Naval Observatory Bulletin 195, March 15, 1962 (unpublished).

Watt, A. D. and R. W. Plush, Power requirements and choice of an optimum frequency for a worldwide standard frequency broadcasting station, J. Research NBS 63D (Radio Prop.), No. 1, 35-43 (July-Aug. 1959).

Watt, A. D., R. W. Plush, W. W. Brown, and A. H. Morgan, Worldwide VLF standard frequency and time signal broadcasting, J. Research NBS, 63D, (Radio Prop), No. 6, 617-627 (Nov.-Dec. 1961).

WWV Standard Frequency Transmissions, Proc. IRE 46, 91 (1958), and subsequent issues.

Appendix 1. NBS Standard Frequency and Time Broadcasting Stations

Station	WWV	WWVH	WWVB	WWVL
Place	Beltsville, Md.	Maui, Hawaii	Boulder, Colo.	Sunset, Colo.
Latitude	38° 59' 33" N.	20° 46' 02" N.	39° 59' N.	40° 2' N.
Longitude	76° 50' 52" W.	156° 27' 42" W.	105° 16' W.	105° 27' W.
Frequency[1]	2.5, 5, 10, 15, 20, 25 Mc/s	5, 10, 15 Mc/s	60 kc/s	20 kc/s.
Offset[1]	-130 pp. in 10^{10}	-130 pp. in 10^{10}	-130 pp. in 10^{10}	-130 pp. in 10^{10}.
Stability[2]	5 pp. in 10^{11}	1 pp. in 10^{10}	2 pp. in 10^{11}	2 pp. in 10^{11}.
Time sig[3]	UT2	UT2	None	None.
Step adj[3]	50 ms	50 ms	None	None.
Time code[4]	3d min of each 5	None	None	None.
Schedule[4]				
Operation	continuous	continuous	continuous	continuous.
Off time	45-49 min past hr	15-19 min past hr 1900-1934 UT	1430-1530 UT	None.
Audio freq[5] Modulation	440/600 c/s	440/600 c/s	None	None.
Schedule	2 min of each 5	3 min of each 5	None	None.
Antennas	Vert. Omnidir.	Vert. Omnidir.	Vert. Omnidir.	Vert. Omnidir.
Type	λ/4 at 2.5 Mc/s λ/2 at other freq	λ/4 at 5 Mc/s λ/2 at other freq	top loaded tuned	top loaded. tuned.
Power radiated	2.5, 20 Mc/s-1 kw 5, 10, 15 Mc/s-10 kw 25 Mc-100 w	5, 10, 15 Mc-2 kw	2 w.	15 w.

[1] The carrier frequencies of WWV and WWVH were offset -150 x 10^{-10} from the United States frequency Standard beginning January 1960; WWVL in April 1960, and WWVB in July 1960. During 1962 the offset will be -130 parts in 10^{10}. This offset enables the time signals, which are locked to the carrier frequency, to maintain close agreement with UT2 time. Corrections to the carrier frequencies as broadcast are available on a weekly basis from National Bureau of Standards, Boulder, Colo., upon request. Corrections for WWV are also published monthly in the Proceedings of the Institute of Radio Engineers.

[2] Frequency adjustments:

WWV	WWVH
As necessary, adjustments of frequency not exceeding one part in 10^{10} are made at 1900 UT. The carrier frequencies are interrupted from 45 to 49 min past each hour.	As necessary, adjustments of frequency not exceeding one part in 10^{9} are made at 1900 UT. The carrier frequencies are interrupted 15 to 19 min past each hour.

[3] Time and time interval adjustments:

Adjustments at WWV of precisely 50 ms may be made in the time signals (second pulses) at the transmitters at 1900 UT when necessary. Such adjustment will be made on the first of the month following the month in which the transmitted time departs from UT2 by more than 50 ms.

WWVH time signals (1 sec pulses) are adjusted, if necessary, each day during the interval 1900 to 1934 UT to be emitted simultaneously with WWV time pulses within ± 1/2 ms.

[4] See appendix 4 for details regarding the time code.

[5] Audio frequency modulations:

The audio frequencies, 440* and 600 c/s, on WWV are transmitted by means of a single upper sideband with full carrier, except on 25 Mc/s. Power output from the sideband transmitter is about 1/3 of the carrier power.

Percent amplitude modulation, double sideband: 440* and 600 c/s signals at 75 percent; voice and sounds pulses, peak is 100 percent at both stations.

Standard Audio Frequencies are broadcast alternately from both WWV and WWVH during two or 3 min of each 5 min interval.

---

* Standard Musical Pitch-A above middle C.

Appendix 2. Standard Time Intervals

WWV time intervals, as transmitted, have the same accuracy as the carrier, ±1 μsec. The frequency offset mentioned above, under Standard Radio Frequencies, applies. Pulses are transmitted at 1 sec intervals. Received pulses have random phase shifts or jitter due to changes in the propagation medium. The magnitude of these changes range from practically zero for the direct ground wave to about 1,000 μsec when received via a changing ionosphere.

Time Signals

Signal schedule: Standard audio frequencies are interrupted at precisely 3 min. before each hour at WWV, and 2 min. before each hour at WWVH. They are resumed exactly on the hour. Except for scheduled silent periods seconds pulses are broadcast continuously except for the 59th pulse of each minute which is omitted. The beginning of a minute is identified by a double pulse consisting of two regular 5 ms pulses spaced by 100 ms. International Morse code announcements of the Universal Time (referenced to the zero meridian) are made each 5 min. from WWV and WWVH. Voice announcements of Eastern Standard Time are made each 5 min. from WWV.

Adjustments of precisely 50 ms may be made in the time pulses when necessary to maintain close agreement with UT2 (see note under Time Intervals).

Corrections, in terms of UT2, of the time signals as finally determined by the U.S. Naval Observatory are published periodically by them.

Radio Propagation Forecasts

A forecast of radio propagation conditions is broadcast in International Morse code from WWV at 19.5 and 49.5 min. after each hour and from WWVH at 9.4 and 39.4 min. after each hour. WWV broadcasts information relating to the North Atlantic radio path and WWVH broadcasts information relating to the North Pacific radio path. Quality is graded in steps ranging from W-1 to N-9 as follows:

W-1 Useless
W-2 Very poor
W-3 Poor
W-4 Poor-to-fair
U-5 Fair
N-6 Fair-to-good
N-7 Good
N-8 Very good
N-9 Excellent

International World Day Service

A symbol indicating the geophysical "state of warning," as declared under the international program of the International Council of Scientific Unions, is broadcast in International Morse Code from WWV at 4.5 and 34.5 min. after each hour and from WWVH at 14.4 and 44.4 min. after each hour.

The following symbols are broadcast to indicate the geophysical conditions:

Symbol	Condition	Remarks
AGI AAAA	Alert	Magnetic Storm with K-index over 5 Outstanding Auroral Display Outstanding increase in Cosmetic Ray flux
AGI --------	Special World Interval in Progress	Geophysical activity of sufficient interest to warrant attention of experimenters throughout the world.
AGI EEEEE	No significant Geophysical events	

Appendix 3. International Coordination of Time and Frequency Transmissions

The United Kingdom and the United States began coordinating their time and frequency transmissions early in 1960. This coordination is the result of an agreement announced by Dr. James H. Wakelin, Jr., Assistant Secretary of the Navy (Research and Development), Dr. Allen V. Astin, Director of the U.S. National Bureau of Standards, and in the United Kingdom by the Astronomer Royal, Royal Greenwich Observatory, and the Director of the National Physical Laboratory.

Coordination was begun to help provide a more uniform system of time and frequency transmissions throughout the world, needed in the solution of many scientific and technical problems in such fields as radio communications, geodesy, and the tracking of artificial satellites.

Participating in the project are the Royal Greenwich Observatory, the National Physical Laboratory, and the Post Office Engineering Department in the United Kingdom, and, in the United States, the U.S. Naval Observatory, the Naval Research Laboratory, and the National Bureau of Standards. This program follows previous cooperative efforts of these agencies to achieve uniformity and simplification in procedures.

The transmitting stations which are included in the coordination plan are GBR and MSF at Rugby, England; NBA, Canal Zone; WWV, Beltsville, Maryland; and WWVH, Maui, Hawaii.

Appendix 4. Time Code on WWV

The timing code provides a standardizing timing basis for use when scientific observations are made simultaneously at widely separated locations. It can be used for example, where signals telemetered from a satellite are recorded along with these pulse-coded time signals; subsequent analysis of the data is then aided by having unambiguous time markers accurate to a thousandth of a second. Astronomical observations may also benefit by the increased timing potential provided by the pulse-coded signals.

This 36-bit, 100-pulse/sec time code, carried on 1,000-c/s modulation, is being broadcast from radio station WWV (2.5,5,10,15,20, and 25 Mc/s). Starting date was January 1, 1961.

1. The code is broadcast for 1-min intervals and 10 times per hour. Except at the beginning of each hour, it immediately follows the standard audio frequencies of 440 c/s and 600 c/s.
2. The code contains time-of-year information (Universal Time) in seconds, minutes, hours, and day of year. It is locked in phase with the frequency and time signals.
3. The code is binary coded decimal (BCD) consisting of nine binary groups each second in the following order: two groups for seconds, two groups for minutes, two groups for hours, and three groups for day of year. Code digit weighting is 1-2-4-8 for each BCD group multiplied by 1, 10, or 100 as the case may be.
4. A complete time frame is 1 sec.
5. The least significant binary group and the least significant binary digit in each group occur first. The binary groups follow the 1-sec reference marker.
6. "On time" occurs at the leading edge of all pulses.
7. The code contains 100-per-second clocking rate, 10-per-second index markers, and 1-per-second reference marker. The 1,000 c/s is locked to the code pulses so that millisecond resolution is easily obtained.
8. The 10-per-second index markers consist of "1" pulses preceding each code group except at the beginning of the second where it is a "0" pulse.
9. The 1-sec reference marker is made up of five "1" pulses followed by a "0" pulse. The second begins at the leading edge of the "0" pulse.
10. The code is a spaced code format; that is, a binary group (BCD) follows each of the 10-per-second index markers. The last index marker is followed by an unused 4-bit group of "0" pulses just preceding the 1 second reference marker.

Session 1. National Bureau of Standards Service to Industry

Paper 1.7. Current Developments in High-Frequency Calibration Services

R. C. Powell*

Calibration services offered by the National Bureau of Standards in the fields of power, voltage impedance, attenuation, and field strength in the frequency range between audio and microwaves are discussed. A brief review is made of services now available and services now being developed. Also considered are plans for the future and problems involved.

Most of the planned improvements are extensions of range and/or accuracy, but part of the discussion covers improvements due to changes in the types of instruments and standards being received for calibration.

1. Introduction

The calibration services to be described are for electronic quantities in the frequency range from 30 kc/s to the practical upper limit of coaxial systems.

The information is divided into four categories: first, the reference standards upon which the calibrations are based and which represent the limit to which a calibration can now be performed, usually at considerable effort and expense; second, the calibration services which are available in the sense that calibrations can be done efficiently but usually to lower accuracies; third, the long-range goals which are given to show what services are being planned. These goals are determined by current needs and practical considerations. Hence, they will change with the state of the art. The fourth category is the immediate future plans for adding new or improved services. These are given collectively for the next five-year period and represent present best estimates based on current funds and personnel limits.

2. High-Frequency Voltage

The national reference standards of high-frequency voltage are thermal converters, a thermister bridge, and an electron beam deflection system. These standards give an estimated error (fig. 1), at 1 v, of 0.1 percent from 30 kc/s to 10 Mc/s, increasing to 0.2 percent at 30 Mc/s, to 1 percent at 100 Mc/s, and to 4 percent at 900 Mc/s.

Calibration services for voltage standards are now provided at 0.03, 0.1, 0.3, 1, 3, and 10 Mc/s over a voltage range of 1 v to 200 v with an uncertainty of about 0.1 percent. This uncertainty increases at higher and lower voltages and higher frequencies. From 0.2 v to 500 v, calibrations can be performed to about 0.2 percent uncertainty at 30 Mc/s and over the same voltage range to about 1 percent at 100 Mc/s. At 300 and 400 Mc/s, services are available from 0.2 v to 20 v with an uncertainty of about 3 percent. Voltages below 0.2 v are obtained using micropotentiometers or attenuators in conjunction with the thermister bridge. From 10 μv to 0.2 v, calibrations can be performed with uncertainties of about 2 percent between 30 kc/s and 10 Mc/s, about 3 percent between 10 and 400 Mc/s, and about 5 percent between 400 and 1000 Mc/s. From 1 μv to 10 μv the uncertainty increases to about 5 percent between 30 kc/s and 400 Mc/s and to about 10 percent between 400 and 500 Mc/s.

The first goal in this area is to provide calibration services to 0.1 percent accuracy from 30 kc/s to 1000 Mc/s over as wide a voltage range as

---

*Chief, HF Impedance Standards Section, Radio Standards Engineering Division, National Bureau of Standards, Boulder Laboratories, Boulder, Colo.

practical. This voltage range is approximately 10 µv to 1000 v. Below 10 µv the uncertainty increases to about 1 percent at 1 µv. The second goal is to provide pulse voltage calibrations, but since research and development in this field have just started, pulse width and peak voltage ranges have not yet been determined.

By the end of the next five years it is expected that services will be available from 30 kc/s to 100 Mc/s with an uncertainty of 0.1 percent between 10 µv and 1000 v, increasing to 5 percent at 0.1 µv. Between 100 and 1000 Mc/s the approximate uncertainties in the calibration services should be 0.1 percent between 0.1 and 20 v, increasing to 1 percent at 10 µv and to 5 percent at 0.1 µv. Only limited pulse service will be available at best.

3. High-Frequency Power

The national reference standards of high-frequency power consist of bolometer bridges and calorimeters. In the power range between 100 µw and 5 w the estimated uncertainty (fig. 2) of these standards is 0.1 percent at 30 kc/s, increasing to 0.25 percent between 30 Mc/s and 300 Mc/s and 0.5 percent between 300 Mc/s and 1 Gc. In the frequency range between 30 kc/s and 1 Gc, higher powers can be measured, but the uncertainty increases to 1 percent at 100 w and to 2 percent at 2 kw. By the use of power dividers, power can be generated from 1 µw to 100 µw at nearly the same accuracy as obtained at the 100 µw level.

Calibration services for power standards are now provided at 0.1, 0.3, 1, 3, 10, 30, 100, 200, 300, and 400 Mc/s between 1 mw and 0.1 w to an accuracy of 1 percent. Below 1 mw, power standards can be calibrated at the above frequencies, but the uncertainty increases to 3 percent at 10 µw and to 30 percent at 1 µw. To calibrate power above 0.1 w, couplers must be used, increasing the estimated uncertainty to 2 percent to 100 w at these frequencies. From 100 to 200 w, calibrations can be performed to 2 percent but only between 0.1 and 30 Mc/s.

The present goals in this area are to provide cw power calibrations with an error of 1 percent from 30 kc/s to the practical upper limits of coaxial systems at practical power levels, to provide pulse power calibrations to about 1 percent accuracy, and to provide similar but more limited services for balanced systems.

By the end of the next five years cw power calibrations at power levels to 10 mw in coaxial systems should be available at fixed frequencies to 10 Gc with approximately the accuracy now obtained at the lower frequencies. Calibration at other than the fixed frequencies should also be available to 10 Gc but with an uncertainty of about 5 percent at best. Limited pulse power services should be available from 0.25 w to 10 kw peak in pulses greater than 1 µsec in duration with an error of about 3 percent.

4. High-Frequency Impedance

The national reference standard of high-frequency impedance is a 1-pf coaxial capacitor. The estimated error (fig. 3) for this capacitance ranges from 0.1 percent at 30 kc/s to 0.5 percent at 300 Mc/s.

Calibration services are now available for resistance, capacitance, inductance, and complex impedance. Expressed in units of impedance, the uncertainty at 30 kc/s is about 0.3 percent from 20 ohms to 100 thousand ohms, increasing to about 1 percent from 1 ohm to 5 meg and to 10 percent from 0.1 ohm to 50 meg. The uncertainty increases with frequency, and at 1 Gc the approximate uncertainty is 2 percent from 20 to 2000 ohms, increasing to 10 percent from about 2 to 50,000 ohms. Three terminal measurements are available below 1 Mc/s to 0.3 percent accuracy over a narrow range of low capacitance.

The present goal in this area is to provide a service for two-terminal impedance standards at 30 kc/s with an uncertainty of 0.1 percent from 100 to 200,000 ohms, increasing to 1 percent from 0.1 ohm to 200 meg. The planned accuracy and range of impedance decreases with frequency. At 1 Gc the goal is to calibrate standards to an uncertainty of about 0.2 percent at 400 ohms, increasing to 1 percent from about 10 ohms to 10,000 ohms. Research on the reference standards is aimed at providing standards over the same impedance range to an order of magnitude less error. While present calibrations do not usually require this accuracy in reference standards, it is anticipated that this is what will be needed for standards using the high-precision connectors now being developed. In both the calibration service and reference standards it is planned to extend the high impedance range an order of magnitude and the low impedance range about one-half an order of magnitude at 1 Gc to about 3 orders of magnitude at 30 kc/s, using three- and four-terminal standards to the same accuracy as the two-terminal standards.

By the end of the next five years it is planned to have services available to the planned range and accuracy at 0.03, 0.1, 0.3, 1, and 3 Mc/s. At 300 Mc/s only the higher impedances will be calibrated to the planned accuracy. Calibrations will be made at other frequencies and magnitudes at about the present accuracies.

5. High-Frequency Attenuation

The national reference standard of high-frequency attenuation is a 30-Mc/s piston attenuator. The estimated uncertainty (fig. 4) of this attenuator is 0.001 db + 0.02 percent over the range 0 to 140 dbs.

Calibration services are now available at 30 Mc/s to 0.005 db + 0.05 percent from 0 to 140 db. At 1, 10, 60, 100, and 300 Mc/s the uncertainty is about 0.05 db + 0.1 percent from 0 to 100 dbs.

Between 0.3 and 6 Gc the uncertainty is about 0.05 db + 1 percent and the range of attenuation from 0 to 50 dbs.

The present goal for the attenuation service is to be able to perform calibrations from 30 kc/s to as high as practical in coaxial systems to an error of not more than 0.005 db + 0.05 percent for attenuation not exceeding 150 db at the highest.

It is planned to have calibration services available from 1 to 300 Mc/s over a range of 0 to 150 dbs. with an uncertainty not exceeding 0.005 db + 0.05 percent by the end of the next five years. By this same time, services should be available over the range 0 to 100 db with an uncertainty of 0.01 db + 0.1 percent between 30 kc/s and 1 Mc/s and also between 300 Mc/s and 10 Gc.

6. High-Frequency Field Strength

Services for field strength consist of calibration of antennas, input attenuators, instrument linearity, and two-terminal voltage. Since the latter three have been discussed, only antenna calibration will be covered here.

The reference standards consist of standard fields produced by calibrated single-turn transmitting loops from 3 c/s to 30 Mc/s or measured by calibrated receiving dipole antennas from 30 to 300 Mc/s. The estimated uncertainty (fig. 5) in the fields produced by the loops is about 3 percent to 5 Mc/s and 5 percent from 5 to 30 Mc/s. The estimated uncertainty of the fields measured by the dipole is about 10 percent to 150 Mc/s and 15 percent between 150 and 300 Mc/s.

Calibration services are now offered for antenna coefficients. The accuracy is mainly dependent upon the reference standards and hence for loops is 3 percent from 3 c/s to 5 Mc/s and 5 percent from 5 to 30 Mc/s. For dipoles the accuracy of calibration is 10 percent from 30 to 150 Mc/s and 15 percent from 150 Mc/s to 300 Mc/s. Calibration accuracy of input attenuators, instrument linearity, and two-terminal voltage is usually dependent upon the instrument itself.

The present goals in this area are to calibrate loop antennas from 3 c/s to 30 Mc/s to 1 percent, to calibrate dipoles from 30 Mc/s to 1000 Mc/s to 5 percent, and to calibrate vertical antennas from 10 kc/s to 30 Mc/s to 3 percent.

Plans for the next five years call for improvement of dipole calibrations from 30 to 1000 Mc/s to an accuracy of 5 percent, to improve loop antenna calibrations from dc to 30 Mc/s to 2 percent, and to provide vertical antenna service to 3 to 5 percent from 10 kc/s to 30 Mc/s.

7. Other

No calibration services are now available in the high-frequency region for noise, power density, interference, or phase shift. Present estimates show that only limited services will be available in the next five years for each of these quantities.

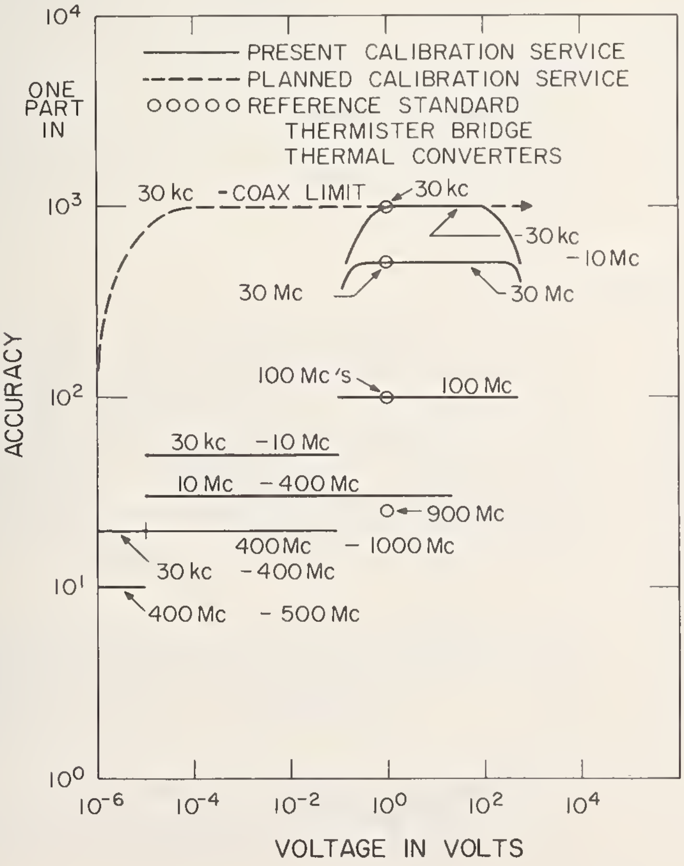


Figure 1. HIGH FREQUENCY VOLTAGE, Accuracy of Voltage Measurements vs Voltage Level.

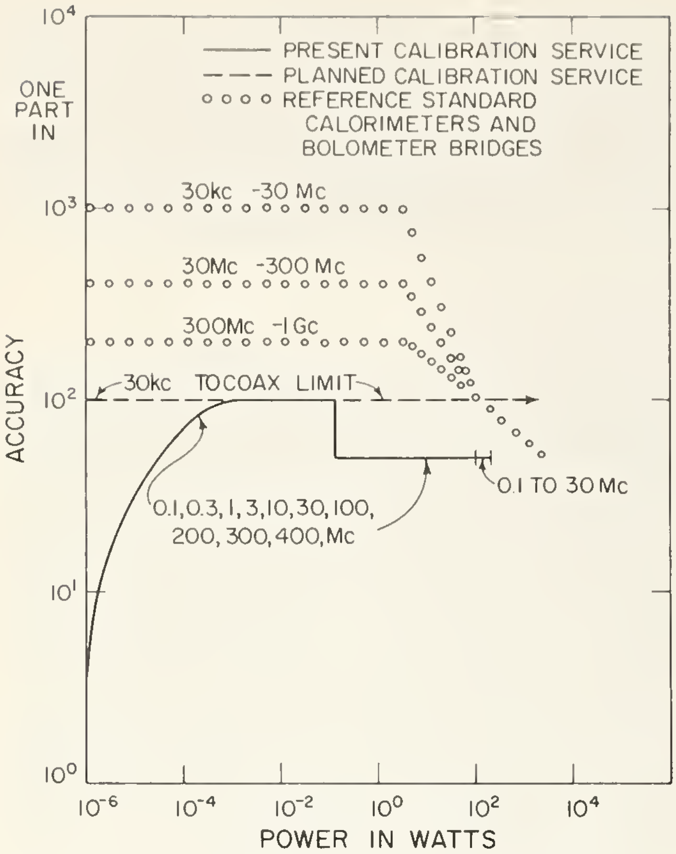


Figure 2. HIGH FREQUENCY POWER, Accuracy of Power Measurements vs Power Level.

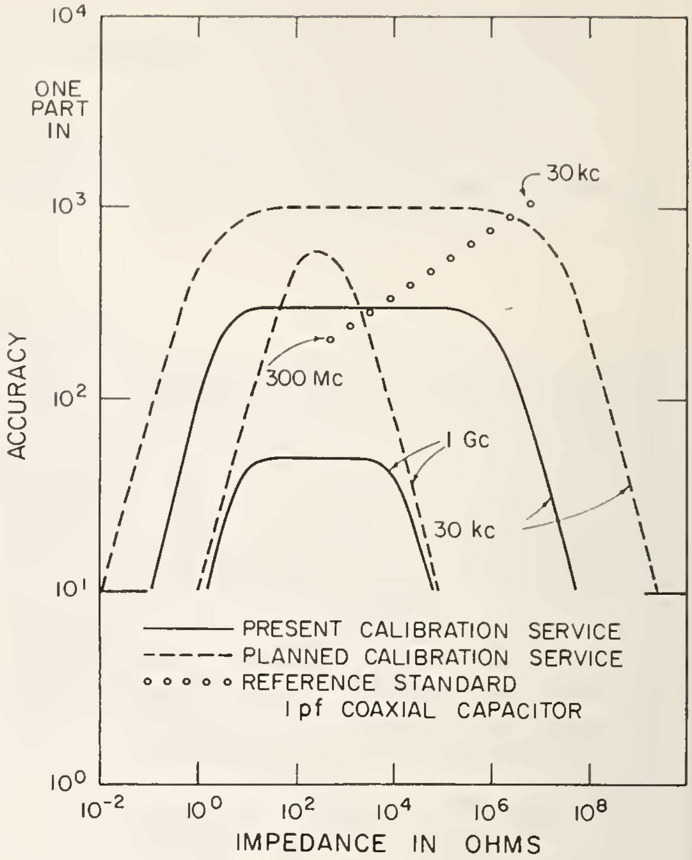


Figure 3. HIGH FREQUENCY IMPEDANCE, Accuracy of Impedance Measurements vs Impedance.

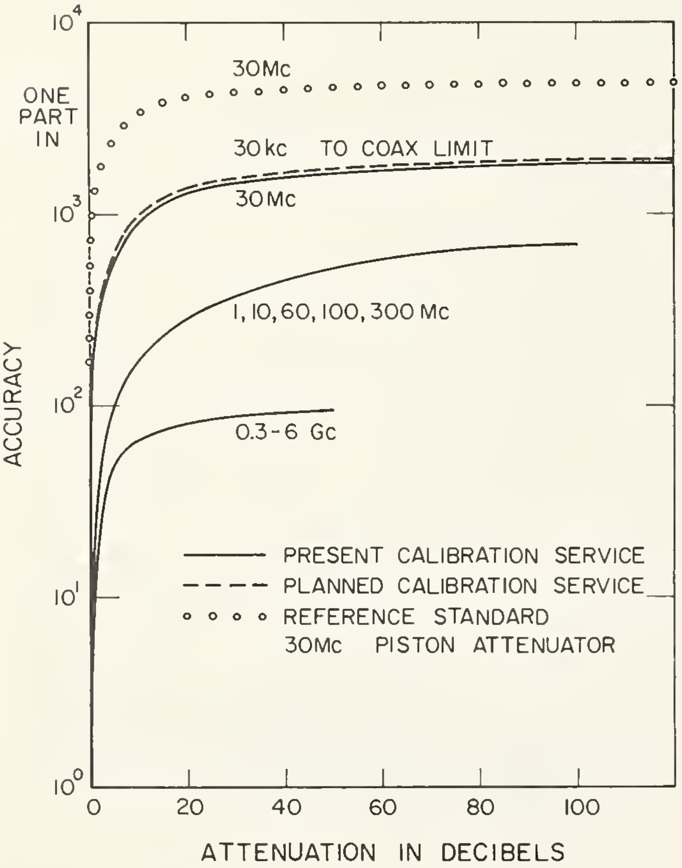


Figure 4. HIGH FREQUENCY ATTENUATION, Accuracy of Attenuation Measurements vs Attenuation.

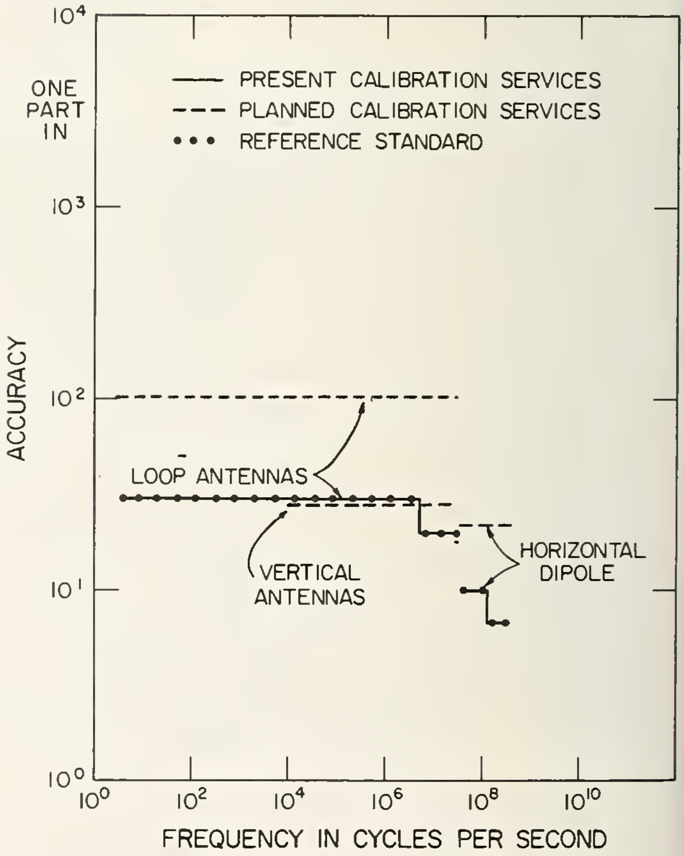


Figure 5. HIGH FREQUENCY FIELD STRENGTH, Accuracy of Antenna Coefficient Measurements vs Frequency.

Session 1. National Bureau of Standards Service to Industry

Paper 1.8. Development of Improved Microwave Calibration Systems

Roy E. Larson*

The development of calibration systems for improved accuracy of measurement and for the extension of frequency coverage and dynamic range is proceeding in the microwave region in the areas of cw power, impedance, attenuation, and noise power. An extension of frequency calibration services is planned to include the measurement of frequency stability of signal sources over a broad range of the spectrum.

New techniques of power measurement developed in the Radio Standards Laboratory will be utilized for calibration services. Improved accuracy of measurement of reflection coefficient magnitude and attenuation will be incorporated in additional calibration systems for extended frequency coverage. Additional noise calibration systems will extend the frequency coverage for this recently initiated calibration service, and an improved accuracy of measurement is expected.

1. Introduction

The development of microwave calibration systems based upon new techniques of measurement and based upon improvements in operating calibration systems is proceeding in the Circuit Standards Division of the National Bureau of Standards Radio Standards Laboratory.

In general, it is hoped that systems can be developed to provide calibration services for the quantities of power, impedance, frequency, attenuation, and noise throughout the frequency range of 2.6 to 40 Gc, with extensions above and below this range in rectangular waveguide as required. This may result in calibration systems being developed to include frequency ranges such as 1.1 to 2.6 Gc and 40 to perhaps 110 Gc in the next few years.

Presently available calibration services in the microwave region have been reported [1].[1] Brief descriptions of the calibration systems utilized to provide these services have been given [2]. Improvements in these systems to obtain better accuracy of measurement or extensions in frequency coverage and dynamic range, in addition to the development of new calibration systems, are presented in the following sections.

2. Power

The development of power measurement techniques in the past few years has provided the basis for the establishment of additional microwave power calibration services at NBS. Calibration systems utilizing the transfer method [3] of comparing power measuring devices are being developed for use in the frequency ranges of 12.4 to 18.0 Gc and 18.0 to 26.5 Gc. This method can provide a measurement of the effective efficiency of a bolometer unit with the advantage that the bolometer unit need not have an extremely low value of reflection coefficient to obtain a good measurement accuracy. From this measurement, further determinations can be made to provide for the calibration factor of the bolometer unit or the calibration factor of a bolometer unit in combination with a directional coupler. The accuracy of power calibrations using this transfer method is expected to be ±2 percent with some of the initial calibration systems to be placed in operation

---

*Chief, MW Calibration Services Section, Circuit Standards Division, National Bureau of Standards, Boulder Laboratories, Boulder, Colo.

[1] Figures in brackets indicate the literature references at the end of this paper.

shortly in the frequency range 12.4 to 18.0 Gc. Further development of this method and the working standards to be used with it should permit an accuracy of ±1 percent. It is hoped that microwave power calibration services also can be made available in the 18.0 to 26.5 Gc frequency range within one year.

Power calibration systems utilizing an impedance technique [4] for the measurement of bolometer unit efficiency are being developed for use in the frequency ranges of 3.95 to 5.85 Gc and 5.85 to 8.2 Gc. The exact procedure for the calibration of interlaboratory standards has not yet been determined, but if bolometric interlaboratory standards are calibrated for efficiency directly with the impedance calibration system, the state of the art indicates that an accuracy of ±0.5 percent is possible. If additional determinations of calibration factor are made following an initial bolometer unit efficiency measurement, the resulting accuracies should be from 1 to 2 percent. The graphical presentation[2] in figure 1 gives the present power calibration capabilities (solid line) and the future power calibration capabilities (dotted line) scheduled for completion within the next year.

3. Impedance

In the area of microwave impedance measurements, progress has been made toward the calibration of reflectors with smaller values of reflection coefficient, and the accuracy of these measurements has been improved. The accuracy of calibration of reflectors with reflection coefficient magnitudes of 0.025 to 0.1 soon will be reported to ±1 percent instead of ±1 1/2 percent. Also, calibrations soon will be performed for reflectors with reflection coefficient magnitudes from approximately 0.001 to 0.025. The accuracy of these measurements will range from ±10 percent for the smaller values to ±1 percent for the larger values.

These improvements are the result of improved designs of precision waveguide sections, improved receiver or detector components, and the development of standard reflectors with small values of reflection coefficient magnitude. Precision waveguide sections have been fabricated with inside dimensional tolerances of 50 μin. Components now are available which will reduce the noise figure and extend the useful linear range of the detection portion of the calibration system.

Concerning the extension of reflection calibration systems to additional frequency ranges, it is expected that calibration systems will be developed and constructed to complete the frequency coverage from 2.6 to 18.0 Gc within the next year. This includes five additional waveguide sizes; the calibration system utilizing WR 284 (2.60 to 3.95 Gc) waveguide probably will be completed earliest.

Figure 2 presents a graphical picture of the present reflection calibration capabilities and the future calibration capabilities expected to be completed within one year.

Work is proceeding on the development of a calibration system for the calibration of two-port phase shifters covering the frequency range 8.2 to 12.4 Gc. It is hoped that an accuracy of ±0.1° may be reported. The calibration system will employ the principals of the modulated subcarrier technique reported by Schafer [5].

4. Frequency

Recent improvements in the calibration of cavity wavemeters includes the use of 5 Mc/s as the initial input for the frequency multiplier chain instead of 100 kc/s. This provides output frequency marker signals with less noise than previously obtained; it reduces the overall multiplication factor and does not deteriorate the precision of the output signal. In using the multiplier chain to calibrate at frequencies between the fixed integral marker output frequencies, a variable-frequency signal is injected into the chain at a point of higher frequency than previously used. This results from improved equipment available, and the method improves the accuracy of the variable output frequency marker obtained in the microwave regions. The availability of signal sources in the higher microwave regions provides greater reliability and ease of operation over the use of harmonic signal generators.

Cavity wavemeters will be calibrated up to 90 Gc in the near future with techniques presently used. An improved frequency-multiplier chain is being designed and constructed using phase-locked oscillators in the microwave region.

The accuracy of calibration of variable cavity wavemeters usually is limited by the dial resettability properties of the device, and with the presently available broadband wavemeters, this limitation usually occurs before a precision of one part in 10^5 is reached. Recent experimental work indicates that present techniques and equipments used to calibrate cavity wavemeters can provide an accuracy of calibration of approximately one part in 10^6. It is hoped that this accuracy capability can be extended to better than one part in 10^7 with the studies now under way.

---

[2] This graphical form represents an attempt to present the three principle parameters needed to describe the basic performance of calibrations using accuracy, frequency, and magnitude of the quantity measured.

The right-hand portion of the graph is one face of a cubical body and is used to plot accuracy versus frequency. The left-hand portion of the graph is an adjacent face of the same cubical body and is used to plot accuracy versus the magnitude of the quantity measured. The left-hand graph is the projection of the plane generated by the line plot on the right-hand graph as the line plot is moved in the direction of increasing magnitude of the quantity measured.

This type of graph permits a convenient display of the calibration information concerning frequency and magnitude of the quantity as related to the accuracy of the measurement.

The calibration of precision signal sources for frequency accuracy and stability has more recently come into demand in the microwave region. Work is proceeding toward the development of improved measurement techniques to accomplish the calibration of precision signal sources in the high frequency and microwave regions.

Figure 3 presents in graphical form the present and future resonance frequency calibration capabilities.

5. Attenuation

The calibration of microwave attenuators at NBS is a service that has been available for some years and that has had extensive use over a broad frequency range. Improvements in this service continue to be made. The calibration systems now in use employ the principles of the IF-substitution technique. The calibration accuracy obtainable from these systems now can be reported to ±0.05 db per 10 db for attenuation difference measurements on variable attenuators and ±0.1 db per 10 db for insertion loss measurements on fixed attenuators. This is accomplished largely by reducing the reflection coefficient of the terminals of the calibration system at the point where the unknown attenuator is inserted. This development already has been incorporated in the frequency range 8.2 to 12.4 Gc (WR 90 waveguide), and the necessary work is proceeding rapidly to extend this accuracy capability to cover the complete frequency range from 2.6 to 26.5 Gc. This will include seven waveguide sizes.

Work is underway to extend the attenuation range over which attenuators can be calibrated with the IF-substitution calibration systems. The principles to be employed are those of the parallel IF-substitution technique. This technique avoids the usual loss through the standard IF attenuator and thus permits the remaining portion of the system to operate over a greater range of attenuation change when comparing rf and IF attenuators.

It is expected that the IF-substitution calibration systems will be extended to cover the frequency range of 26.5 to 40 Gc (WR 28 waveguide) during the next year. This extension in frequency would accomplish for attenuation calibration services one of the primary objectives stated earlier.

An attenuation calibration system has been developed and constructed following the principles of the modulated subcarrier technique. [6] This system, utilizing WR 90 waveguide (8.2 to 12.4 Gc), has been used to date to perform some special calibrations both internal and external to NBS. The system will be developed further with the purpose of providing calibration services of attenuation measurements to accuracies of approximately ±0.001 db over a limited range of attenuation values. In addition to this system, it is planned to develop and construct a similar calibration system utilizing WR 62 waveguide (12.4 to 18.0 Gc) during the year.

Figure 4 is a graphical representation of the present and future microwave attenuation calibration capabilities.

6. Noise

The newly initiated microwave noise calibration service represents a long-awaited fulfillment of a need for measurements in the microwave region. The calibration system and the National Reference Standard was developed for use in WR 90 waveguide (8.2 to 12.4 Gc), and calibration service presently is provided at 9.0, 9.8, and 11.2 Gc in this frequency range. As in all of the microwave calibration services, calibration of microwave noise sources will be made available throughout the frequency range 8.2 to 12.4 Gc. Initial calibrations requested have been for argon gas-discharge noise sources with an excess noise ratio of approximately 15.6 db. In reporting upon calibrations performed, the effective noise temperature and the excess noise ratio are given to accuracies of ±250 °K and ±0.1 db, respectively.

Some additional developments already have revealed that an improved accuracy for noise calibrations can be accomplished in the near future. With improved precision rf attenuators and improved calibration of them, with newly available components of lower noise figure for the detection circuit, and with additional rearrangement of the radiometer circuit, an accuracy of the measurement of excess noise ratio of ±0.06 db or better can be obtained.

It is expected that the microwave noise calibration service can be extended to include the frequency range 12.4 to 18.0 Gc in the next year. It is hoped that the frequency range 2.6 to 3.95 Gc also can be completed during this year. This planned work includes the development and construction of suitable hot-body reference standards for these two frequency ranges.

Figure 5 presents in graphical form the present and future microwave noise calibration capabilities.

7. Conclusion

This survey has given some of the present developments toward improved microwave calibration services and has included some of the work expected to be completed within the next year. It is recognized that there still will be limitations in frequency range, magnitude, and accuracy in calibration services for the quantities mentioned. Also, the development of measurement techniques and calibration systems for additional quantities is urgently needed. Future work is planned to overcome these limitations as soon as possible.

In considering the remarks made about better calibration systems and higher accuracy measurements in the microwave region, one point should be mentioned. With improvements in measurement techniques producing higher measurement accuracies, greater demands are made upon the quality and capabilities of interlaboratory standards. If higher accuracies are to be transferred from one place to another by means of an interlaboratory standard, this standard logically must display characteristics of high precision and good stability.

In the microwave region, where most devices used as interlaboratory standards are constructed with rectangular waveguide and the corresponding terminals, the components presently available sometimes fall short of the desired precision and stability. It is believed this, in part, is due to the continued practice of manufacturing waveguide components to cover a broad band of frequencies. The broadband components of course, are usually suitable for most laboratory-type work. But it would not be undesirable to place more emphasis upon the design of waveguide components specifically tailored to meet the coming needs of the standards laboratory.

8. References

[1] Federal Register **26**, No. 231, Part 201 - Electricity, high frequency and microwave regions.

[2] Larson, R. E., Microwave measurements in the NBS Electronic Calibration Center, Proc. IRE **109**, Pt. B, Suppl. 23, 644-650 (1962).

[3] Engen, G. F., A transfer instrument for the intercomparison of microwave power meters, IRE Trans. Instr., **I-9**, No. 2, 202 (Sept. 1960).

[4] Engen, G. F., A bolometer mount efficiency measurement technique, J. of Research, NBS **65C** (Eng. and Instr.), No. 2, 113-124 (Apr.-June 1961).

[5] Schafer, G. E., A modulated subcarrier technique of measuring microwave phase shifts, IRE Trans. Instr. **I-9**, No. 2, 217 (Sept. 1960).

[6] Schafer, G. E., Bowman, R. R., A modulated subcarrier technique of measuring microwave attenuation, Proc. IRE **109**, Pt. B, Suppl. 23, 783-786 (1962).

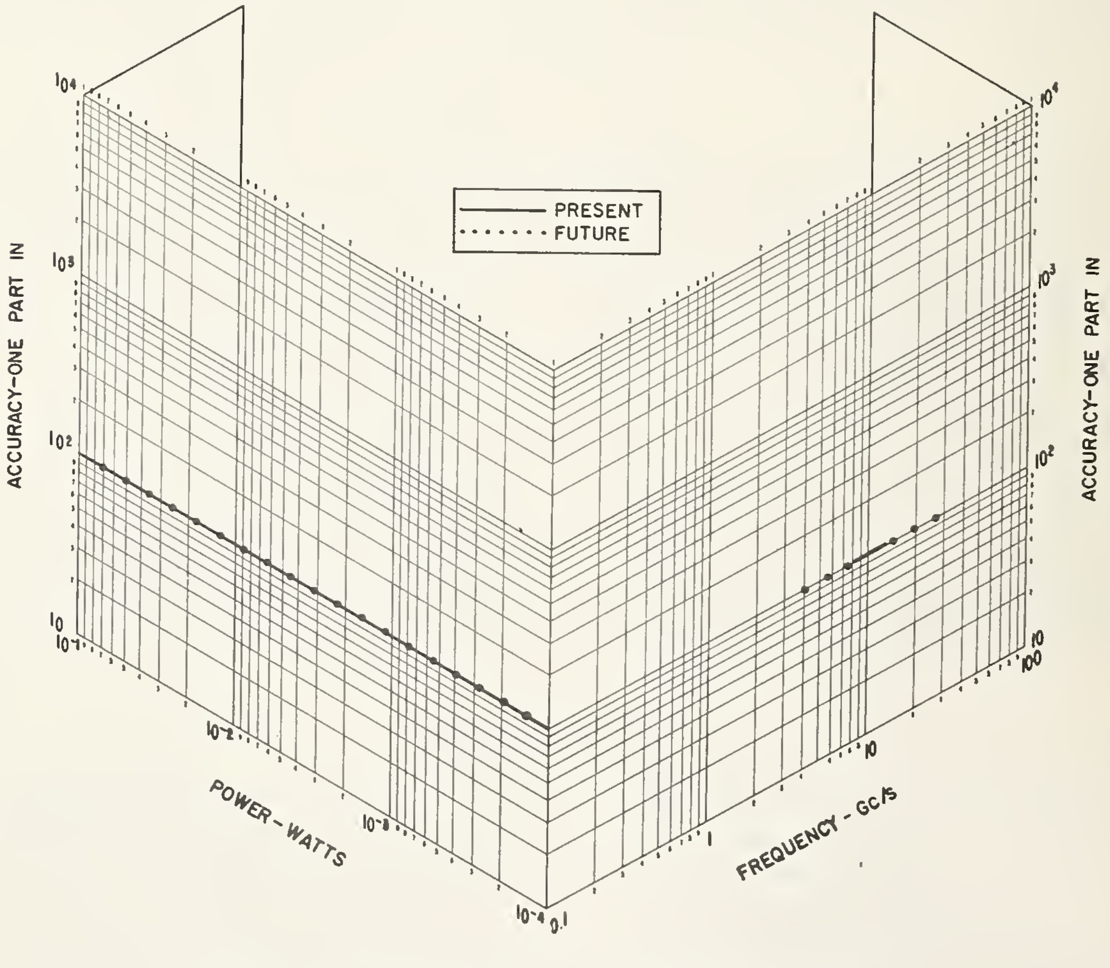


Figure 1. Microwave Power Calibration Services.

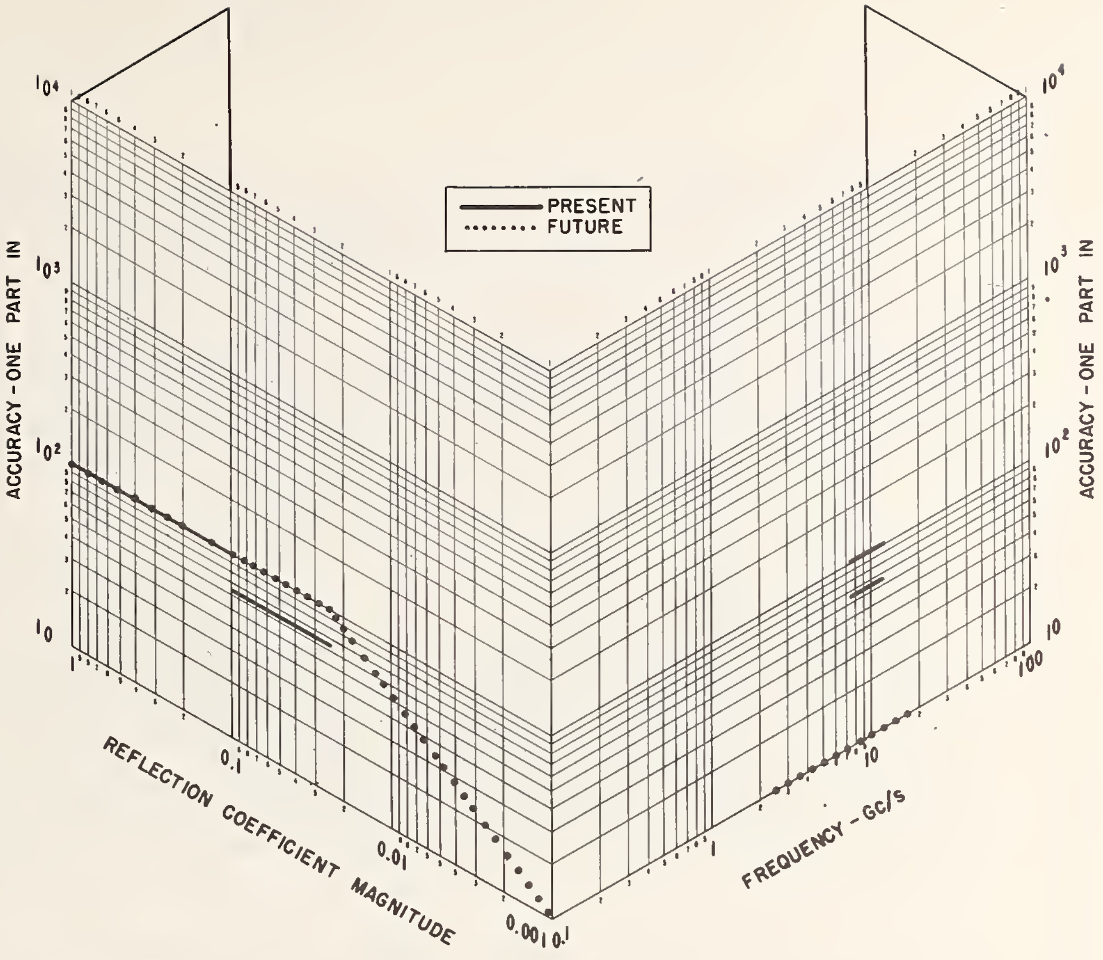


Figure 2. Micro wave Reflection Calibration Services.

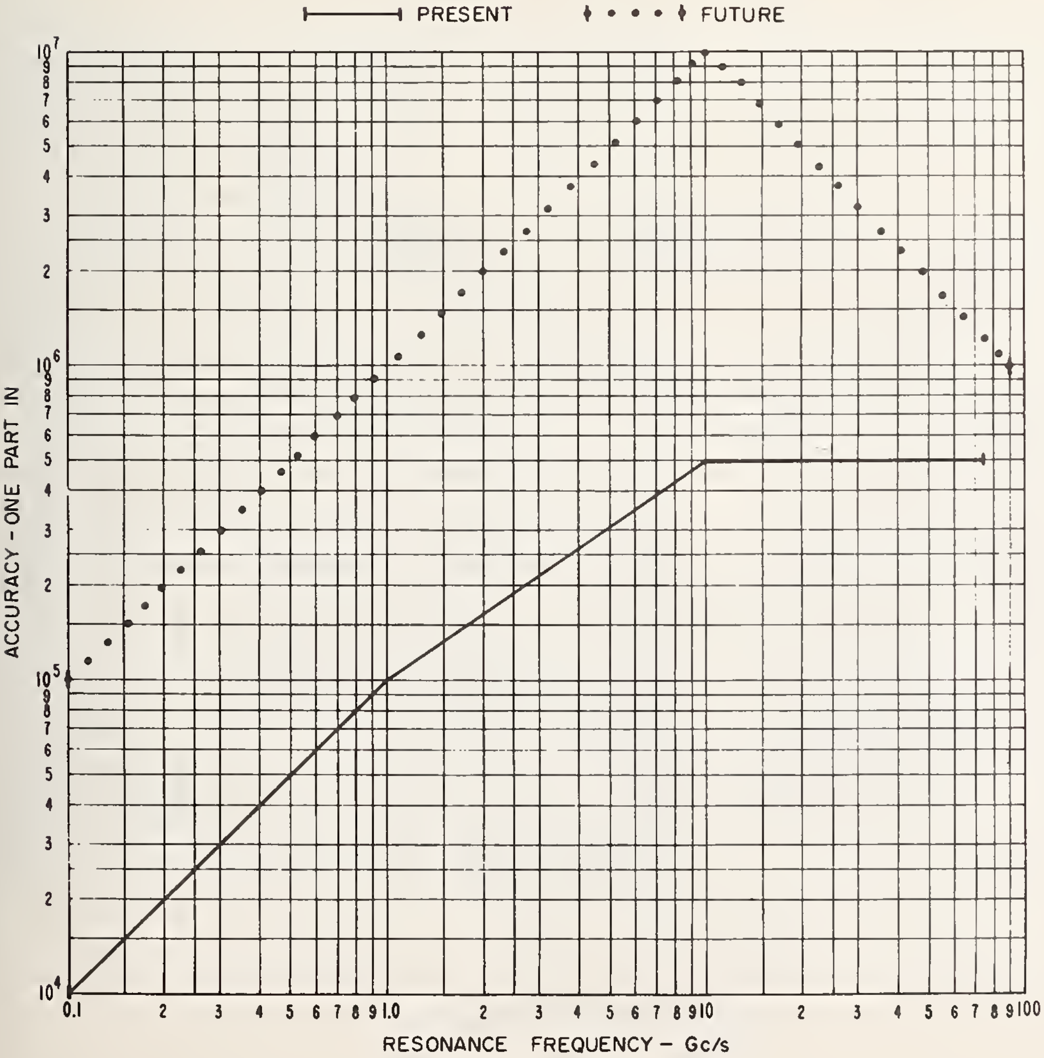


Figure 3. Frequency Calibration Services.

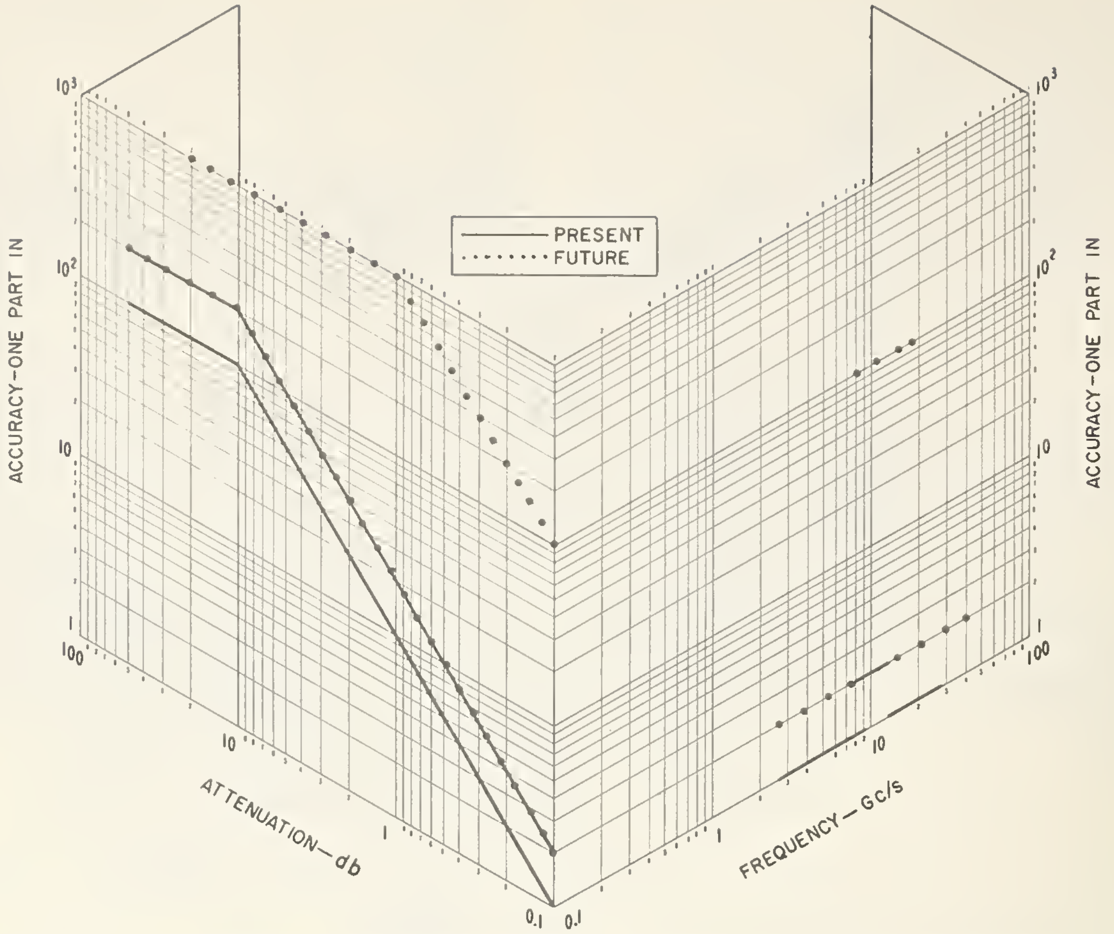


Figure 4. Microwave Attenuation Calibration Services (Variable Attenuators).

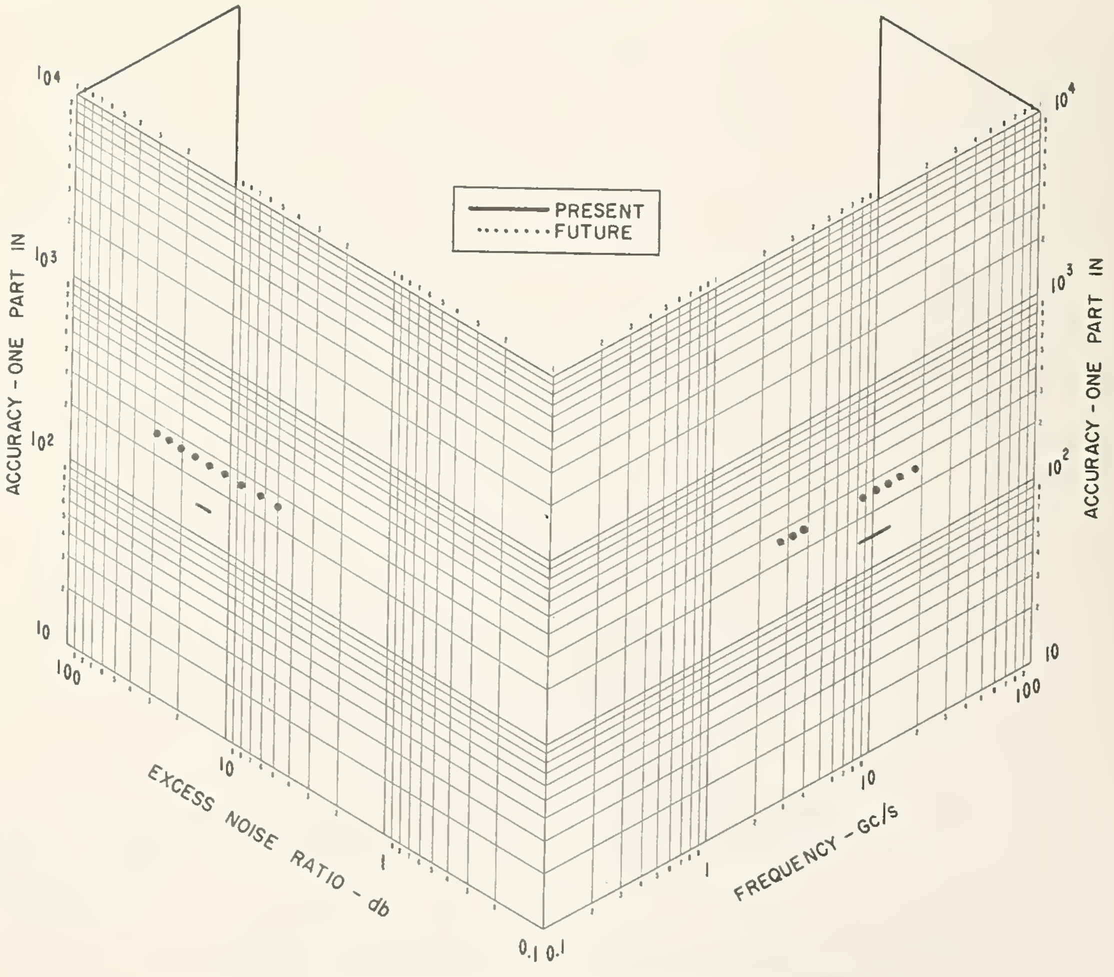


Figure 5. Microwave Noise Calibration Services.

Session 1. National Bureau of Standards Service to Industry

Paper 1.9. Measurements and Standards for Radio and Microwave Materials at the National Bureau of Standards

John L. Dalke*

The science of materials, particularly as it relates to electromagnetism, has evolved primarily from the activities of two different groups, engineers and solid state physicists. The need for a consensus about methods and terminology has become acute. It is a basic function of National Bureau of Standards to promote this consensus through research and publication as well as calibration services for several electromagnetic properties of materials to be utilized by other laboratories for standardization purposes. Facilities are available for determining dielectric, conductivity, and magnetic properties over the radio and microwave frequency ranges. Examples of how some measurement problems have been met at the National Bureau of Standards as well as a list of ranges and accuracies are discussed.

1. Measurements and Standards

The electromagnetic material parameters selected for measurement are those which are basic to both science and engineering. Methods and techniques for achieving interaction between the electromagnetic wave and the specimen are of prime importance. The geometry of the specimen and its relationship to the rest of the electromagnetic structure and the means of launching the signal in the material are paramount. Many arrangements are well known; however, their limitations and the importance of accuracy are not always as well known. Examples of how some of these problems have been met at NBS are described briefly.

There is no question that standardization in the area of electromagnetic material measurements is essential. "Round robins" conducted by organizations such as the IRE and the ASTM as well as individual observations and NBS experience confirm this need. Further support for this need is provided by the Aerospace Industries Association Measurement Research Conference, Meetings No. 19 and 20[1][1] held in February 1962. The main conclusion of these conferences was the need for reference specimens and the round robin technique for comparing measurements in industry. NBS facilities are designed to calibrate standard reference specimens submitted by others. For this purpose, specimens of known stability and character are highly desirable. NBS also has a program directed towards characterizing selected materials for use as standard references. The parameters selected for standardization are the three constitutive electromagnetic parameters. They are the dielectric constant or permittivity, the magnetic permeability and the conductivity. Each of these has a complex frequency spectra. The complex parameters have real and imaginary components. The real part represents the energy storage mechanism, while the imaginary part represents the loss mechanism. In addition, the dispersion is often a function of quantities such as temperature, stress, superimposed states, electric or magnetic fields and time. For materials with anisotropies, either inducted or inherent, the constitutive parameters become tensors. The components of these tensors are also usually complex quantities having real and imaginary parts. All these factors make a variety of measuring methods necessary. Careful attention must also be given to environmental detail which provides superimposed conditions that may affect the results. In some cases the environment is imposed to

---

[1] Figures in brackets indicate the literature references at the end of this paper.

*Chief, Radio and Microwave Materials Section, Radio Physics Division, National Bureau of Standards, Boulder Laboratories, Boulder, Colo.

produce specific material effects such as Faraday rotation while in others factors such as temperature or humidity are always present and their effects must be known. Representative NBS measurement methods related to determining material characteristics appear in the following paragraphs.

NBS radio frequency standards for permittivity and permeability measurements are based on precise variable length coaxial lines providing calculable incremental capacitances and incremental inductances for which significant circuit parameters such as skin depth, fringing capacitance, and lead inductance are taken into consideration [2, 3]. These devices provide the necessary reference standards for calibrating precision "electrode systems" used in a variety of structures and techniques employed in resonance and bridge measurements at radio frequencies. Electrode systems are arrangements for incorporating the specimen into the electromagnetic circuit in a known way. Perturbation techniques are employed. This entails making measurements with the specimen in and out of the circuit. Associated with this a length change is involved. The length change is calibrated in terms of an appropriate electromagnetic characteristic of the circuit such as capacitance or inductance. Up to about 500 Mc/s precise plane parallel circular electrodes driven by a precision micrometer calibrated in terms of the capacitance between the electrodes are used for measuring complex permittivity. The permittivity is calculated from differences in the micrometer readings with the specimen in and out. The disc-shaped specimen is no larger than the electrode diameter. It may be smaller. With this technique, fringing capacitance produces only second order effects. See figure 1 for a photograph of representative electrode systems of this type. Foil electrodes are applied to the specimen for measurements in these holders. Similarly the permeability is obtained up to several thousand megacycles on toroidal specimens from precision line length changes. Bridge and Q-meter techniques are used to several megacycles. Above this, the resonance or cavity technique is used most frequently although some measurements are made using standing wave machines. The cavities are coaxial lines, reentrant capacitatively loaded coaxial structures, or microwave cavities with modes selected to provide high accuracy. The loss tangent is determined from the shape of the resonance curve or from the resistive components of the bridge when the bridge method is used. At the higher or microwave frequencies, the preferred specimen geometries are rods and spheres.

Among auxiliary radio techniques used at NBS is an RF permeameter [4, 5] which provides complex permeability measurements on toroids without windings (fig. 2). It operates in the difficult frequency range from a few hundred kilocycles to about fifty megacycles. An equally interesting instrument also developed at NBS is the permittimeter [6] in which a dielectric or semiconductor in the shape of a toroid has an electrical field applied tangentially to the torous (fig. 3). The utility of this device is that it allows complex dielectric constant or complex conductivity measurements without applying electrodes to the specimen. It is usable from a few hundred kilocycles to several megacycles. However, since it utilizes impedance transformation, the dielectric constant must be substantial to achieve sufficient accuracy on available bridges and resonance devices. It is particularly advantageous for materials having a large dielectric constant such as ferrites.

For precise complex permeability measurement on low loss and low-permeability materials, a special low-impedance low-frequency Maxwell bridge [7] has been developed. To achieve high accuracy, the holder is a coaxial line whose center conductor terminal is placed in the high-impedance arms of the bridge. The toroidal specimen is introduced into the coaxial holder by sliding the toroidal specimen over the center conductor via terminal openings in the detector and high-impedance arms adjacent to the center conductor (see fig. 4). Also of interest is a technique for reducing errors in permeability measurements using coils on toroids [8]. With this approach, coils with a small number of turns have provided adequate accuracy employing convenient inductance measurements.

In addition to complex permeability, equipment has been developed for measuring static and dynamic magnetostriction. The static technique uses a Fabry-Perot optical interferometer with a sensitivity approaching five parts in 10^8. Limited total loss measurements using a calorimetric technique operating a nominal frequency of 300 kc/s are also made. A quasistatic automatic flux loop plotter to supply supplementary magnetic information has also been developed.

A basic magnetic parameter important in both applications and theory is saturation or spontaneous magnetization. A vibrating sample high-precision magnetometer [9] and an associated self-calibrating technique have been developed to meet this need. The magnetometer described in reference 9 vibrates the specimen along the axis of the sensing coils. A more recent modification of the NBS version of this magnetometer vibrates the specimen perpendicular to this direction and still provides adequate sensitivity.

A microwave standard for measurement of surface impedance, skin depth and conductivity has been developed (fig. 5) [10]. As a result, a calibration service for microwave conductivity is available. Q's approaching 96 percent of the theoretical value have been obtained and the departure from 100 percent is fairly well understood. At lower frequencies, conductivities are determined using a four-electrode technique or the permittimeter mentioned earlier.

The best microwave dielectric measurements are made using a TE_{011} circular cylindrical resonator [11] with a rod-shaped specimen on the axis (fig. 6). A perturbation technique is used. Curves and tables have been made to be used in evaluating the results. The calculations may be performed in a few minutes.

Also of interest is work done on errors in dielectric measurements caused by a sample inserted in a hole in a cavity [12]. Improved accuracy is possible by applying such corrections.

At microwave frequencies, the magnetic properties of materials such as ferrites are tensor quantities. An exact solution for a cylindrical gyromagnetic material rod in a degenerate cavity [13] has been obtained which provides the components of tensor permeabilities. Facilities have been developed for making complex tensor permeability measurements at S and X band frequencies. Associated linewidth and g-factor [14] measurements are also made at L, S, and X bands.

Temperature is a significant parameter in material evaluation both from a practical and theoretical viewpoint. Some limited temperature dielectric measurements can be made approaching 800 °C at selected frequencies.

2. Services, Frequency Ranges and Accuracies

NBSBL officers calibration services for standard reference specimens of magnetic, dielectric or conductivity materials at radio and microwave frequencies. Since spontaneous or saturation magnetization is an important parameter in microwave applications, such measurements are also made, although they are not considered a regular calibration service. Similarly special measurements on magnetostriction, hysteresis curves, reversible permeability, temperature coefficients of permeability or permittivity, Curie points, etc., can be made depending on the urgency of the demand. Table 1 lists representative types of measurements, ranges, and order of magnitude of accuracy for radio material calibrations at NBSBL. The accuracy is a function of the frequency, the magnitude of the complex material quantity and to what extent the complex components interact. The accuracy given in the table usually represents the poorest accuracy attained for the more conventional magnitudes of the material's parameters. In most cases, tha accuracy achieved for a given measurement is as good as or better than the value in the table depending upon the material and frequency involved.

In selecting reference specimens, emphasis must be placed on temperature, humidity, and long-term aging stability. The specimen should be independent of these three environmental control parameters, or it should be known that the specimen provides reproducible results when returned to a given temperature and moisture content (either internally and/or on the surface). The aging should be negligible over the period of time for which it is to serve as a standard reference including the time from the original calibration to recalibrations. If aging is significant and the pattern is known, the specimen may still be usable in some instances. Preferably, the specimens selected for references should be well-characterized known compositions. If knowing the composition is not practical, at least knowing the behavior history is highly desirable. NBS calibrates specimens submitted by the user, and, in some cases, can supply the user with standard reference specimens.

TABLE 1

Quantity	Methods of measurement	Range	Nominal accuracy
I. Conductivity:			
1. Complex conductivity	Four-terminal radio transformer bridge; permittimeter with bridge and Q meter; TE_{011} mode cavity.	30 kc/s; 100 kc/s to 1 Mc/s; X band.	$\sigma' \sim 2\%$; tan $\delta \sim 5$ to 10%; (apparent skin-depth for X band $\sim$ 2%).
2. Complex tensor conductivity	Degenerate cavity.	No service currently available. Some development work at X and K bands has been done.	
II. Dielectrics:			
1. Complex dielectric constant or complex dielectric susceptibility	Precision electrode systems with bridges, Q-meters, and re-entrant cavities; standing wave machines, precision cavities.	30 kc/s to 10Gc; 30 Gc and 100 Gc under development.	$\varepsilon' - 1 \sim 0.1\%$ to 10% depending on $\lvert \varepsilon^* \rvert$ type of material and frequency; tan $\delta \sim$ 5% to 10%.
III. Magnetics:			
1. Complex permeability			
a. Initial permeability	(1) Special designed bridge. (2) Permeameters or demountable coils with Q-meters or bridges.	Data provided from 30 kc/s to 100° Mc/s.	$\mu' \sim 1\%$; tan δ $\sim 10\%$

TABLE 1.--Continued

Quantity	Methods of measurement	Range	Nominal accuracy
III. Magnetics--Continued			
	(3) Variable length reentraint cavities. (4) Variable length half wave coaxial cavities.		
b. Reversible permeability	Permeameters, reentrant cavities, half wavelength coaxial cavities.	Data provided from 50 kc/s to 1000 Mc/s. Fields corresponding to currents up to 100 amp turns applied parallel to RF field.	$\mu' \sim 1\%$; tan $\delta \sim 10\%$.
c. Temperature coefficient of Permeability	Permeameter	Frequency = 1 Mc/s temperature range = 20 °C to 250 °C.	Same.
2. Ferrimagnetic resonance			
a. Kittel resonance Linewidth and $\underline{g}$ factor	Cavity perturbation technique using spherical samples in TE_{nom} cavities.	Measurements provided at L-S-X band frequencies.	$\underline{g} \sim 1\%$; $\Delta H \sim 5\%$.
b. Tensor permeability	Cavity methods using rod shaped samples in TM_{110} cavities.	Measurements at S and X band frequencies as a function of d-c field.	Varies with magnitude of permeability; $\underline{g} \sim 1\%$; $\Delta H \sim 5\%$.
3. Saturation magnetization	Vibrating sample magnetometer.	M may be given as a function of applied field up to 10,000 oersteds.	$\sim 1\%$.
4. Hysteresis loops	Quasistatic automatic flux loop plotter.	Flux densities B to 50 kilogauss.	B to $\sim 1\%$ for sufficiently high flux densities.
5. Dynamic magnetoelastic properties	Coil measurements on toroids using inductance bridges.	Data usually given as function of d-c field.	Varies with particular quantity measured.
6. Total loss measurements	Calorimeter	Data provided at 300 kc/s	$\sim 2\%$.
7. Curie temperature	Force magnetometer	Measurements up to 600 °C	± 5 °C.

$\varepsilon * = \varepsilon' - \underline{j}\varepsilon''$ $\mu * = \mu' - \underline{j}\mu''$

$\sigma * = \sigma' - \underline{j}\sigma''$ $\tan \delta = \frac{\mu''}{\mu'}$, etc.

3. References

[1] Aerospace Industries Association Calibration and Measurement Project Report of Spring 1962 Meeting, May 3-4, 1962, Project Sponsor, Sperry Gyroscope Co., Great Neck, New York.

[2] Powell, R. C., Jickling, R. M., Hess, A. E., High frequency impedance standards at the National Bureau of Standards, IRE Trans. Instr. **I-7**, Nos. 3 & 4, 270 (Dec. 1958), (The incremental capacitor described is the standard reference for calibrating dilectric electrode systems.)

[3] Harrington, R. D., Cavity Techniques for permeability measurements in the VHF region, Proc. Electronic Components Conf., May 1955, (Note the variable length coaxial line which serves as a standard.)

[4] Rasmussen, A. L., Hess, A. E., R-F permeameter techniques for testing ferrite cores, Elec. Mfg. 61, No. 5, 86 (May 1958).

[5] Hoer, C. A., Rasmussen, A. L., Equations for the radiofrequency magnetic permeameter, J. Research NBS. **67C** (Eng. and Instr.), No. 1, 69-76 (Jan.-Mar. 1963).

[6] Powell, R. C., Rasmussen, A. L., A radiofrequency permittimeter, IRE Trans. Instr. **I-9**, No. 2, 179-184 (Sept. 1960).

[7] Rasmussen, A. L., Powell, R. C., A low-impedance Maxwell bridge for measuring toroidal magnetic materials from 1 kc to 100 kc, Proc. IRE **50**, No. 12, 2505-2506 (Dec. 1962).

[8] Danielson, B. L., Harrington, R. D., A technique for reducing errors in permeability measurements with coils, Proc. IRE **48**, No. 3, 365 (Mar. 1960).

[9] Frederick, N. V., A vibrating sample magnetometer, IRE Trans. Instr. **I-9**, No. 2, 194-196 (Sept. 1960); Frederick, N. V., A simple calibration technique for vibrating sample and coil magnetometers, Proc. IRE **49**, No. 9, 1449 (Sept. 1961).

[10] Bussey, H. E., Standards and measurements of microwave surface impedance skin depth, conductivity and Q, IRE Trans. Instr. **I-9**, No. 2, 171-175 (Sept. 1960).

[11] Bussey, H. E., Cavity resonator dielectric measurements on rod samples, Insulation, p. 26 (Nov. 1959); 1959 Annual Report, Conf. Electrical Insulation, Natl. Res. Council-Natl. Acad. Sci., Wash., D.C., Jan. 1960; Javitz, A. E., Research progress in dielectrics, 1959, Elec. Mfg. 65, No. 1, 60-70 (Jan. 1960).

[12] Estin, A. J., Bussey, H. E., Errors in dielectric measurement due to a sample insertion hole in a cavity, IRE Trans. Microwave Theory and Tech. **MTT-8**, No. 6, 650-653 (Nov. 1960).

[13] Bussey, H. E., Steinert, L. A., An exact solution for a cylindrical cavity containing a gyromagnetic material, Proc. IRE **45**, No. 5, 693 (May 1959); Bussey, H. E., Steinert, L. A., Exact solution for a gyromagnetic sample and measurements on a ferrite, IRE Trans. Microwave Theory and Tech. **MTT-6**, No. 1, 72 (Jan. 1958).

[14] Preston, C. C., Case, W. E., Tables to facilitate the determination of the ferrimagnetic resonance linewidth of nonmetallic magnetic materials, NBS Tech. Note 173 (Feb. 1963).

Figure 1. Representative Dielectric Electrode Systems Used at NBSBL. (Tall structure to the right is a coaxial variable length re-entrant cavity).

Figure 2. RF Permeameter for Measurements on Toroidal Ferrimagnetic Specimens. (This holder avoids windings on the specimen. Typical magnetic specimens are at the lower right. The coil visible at the bottom of the coaxial structure produces a circular magnetic field about the center conductor. The toroidal specimen is inserted in this field when measurements are made. The terminals project from the bottom of the structure).

Figure 3. RF Permittimeter. (The walls, end plates and center conductor of the coaxial structure are made of ferrite. A primary winding excites the circular electric field about the center post. Toroidal dielectric or conductivity substances are inserted in this field for evaluation. A toroidal plastic container with a channel in which liquids can be placed provides measurements on the dielectric or conductivity properties of fluid substances).

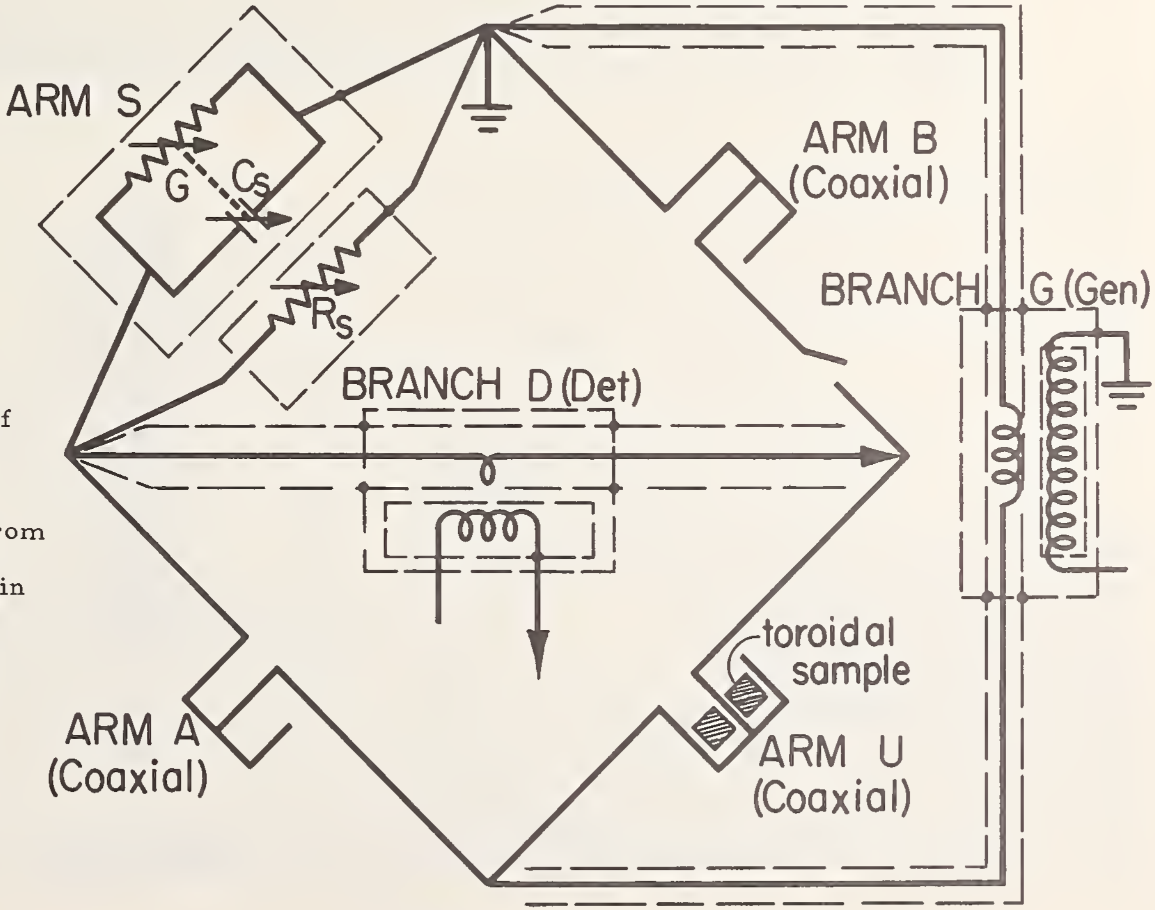


Figure 4. Schematic of Low Impedance Maxwell Bridge for Measuring Toroidal Magnetic Materials from 1 kc to 100 kc. (The specimen is inserted in the holder by opening connections in the detector and Arm B. Much small complex permeabilities are measurable with this instrument than with more conventional techniques).

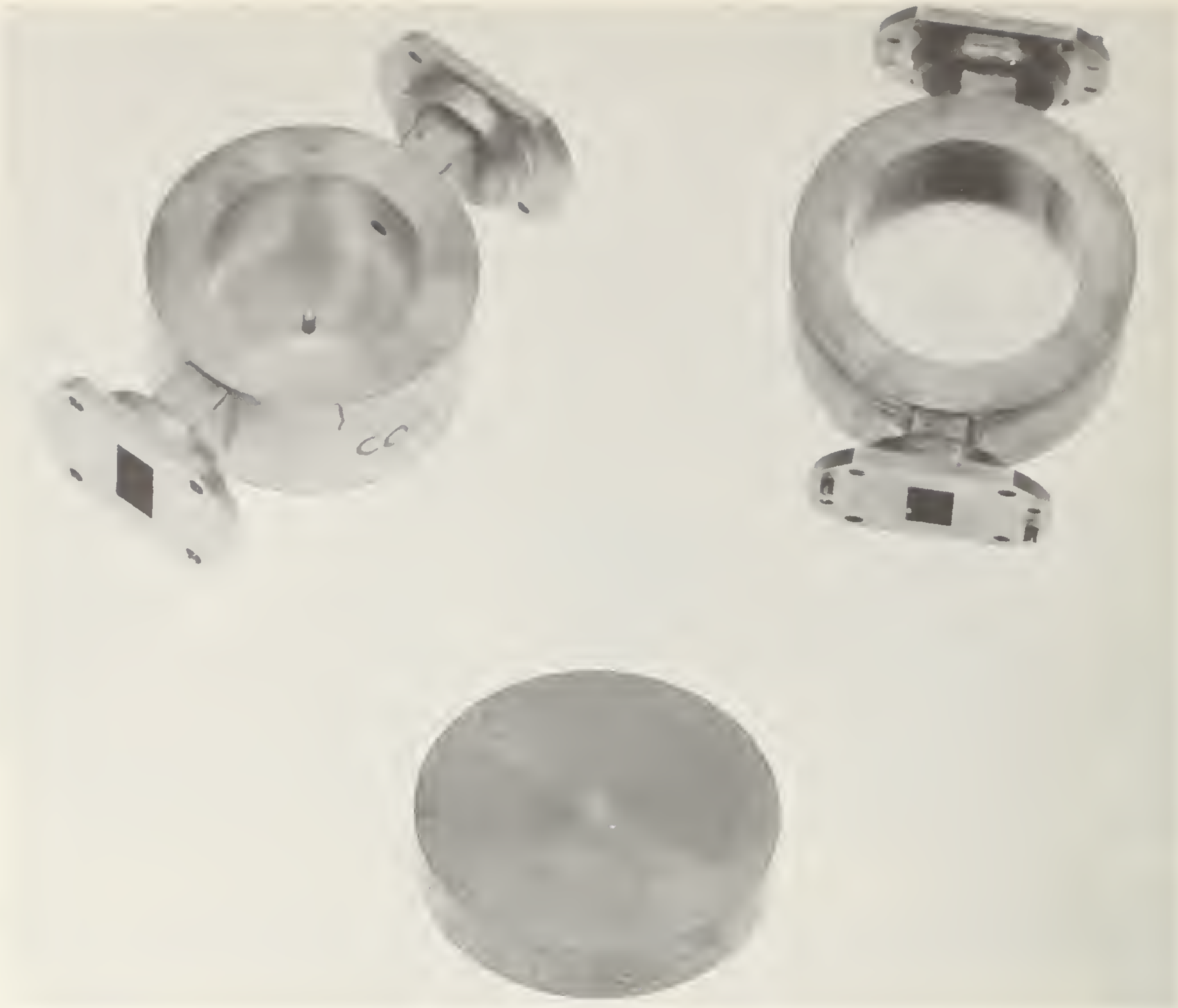


Figure 5. TE_{011} Mode Cavity Resonators at 9200 Mc and a Flat Eng Plate to be Evaluated. (Cavity with post is used for standardization of microwave surface impedance, skin depth, conductivity and Q. The post suppresses unwanted modes).

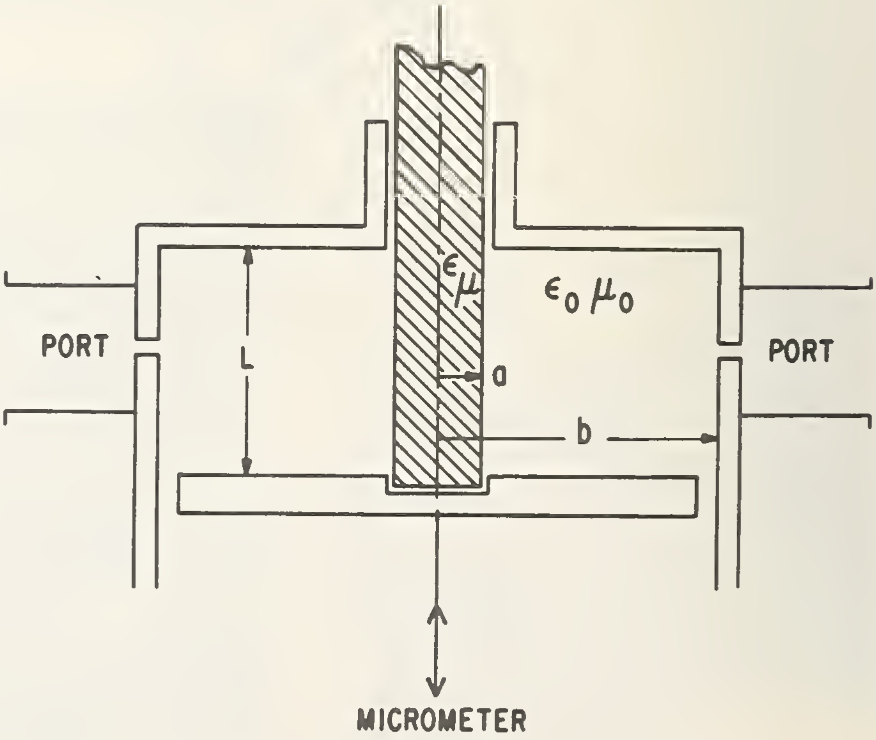


Figure 6. Schematic Arrangement of a Variable Length TE_{011} Mode Cavity at 9200 Mc used for Making Complex Dielectric Constant Measurements. (The actual cavity is supplied with mode suppressing identations).

NBS Miscellaneous Publication 248
Proceedings of the 1962 Standards Laboratory Conference

Session 2. Error Analysis of Measurement Systems

Paper 2.1. Realistic Evaluation of the Precision and Accuracy of Instrument Calibration Systems *

Churchill Eisenhart * *

Calibration of instruments and standards is a refined form of measurement. Measurement of some property of a thing is an operation that yields as an end result a number that indicates how much of the property the thing has. Measurement is ordinarily a repeatable operation, so that it is appropriate to regard measurement as a production process, the "product" being the numbers, i.e., the measurements, that it yields; and to apply to measurement processes in the laboratory the concepts and techniques of statistical process control that have proved so useful in the quality control of industrial production.

Viewed thus it becomes evident that a particular measurement operation cannot be regarded as constituting a measurement process unless statistical stability of the type known as a state of statistical control has been attained. In order to determine whether a particular measurement operation is, or is not, in a state of statistical control it is necessary to be definite on what variations of procedure, apparatus, environmental conditions, observers, operators, etc., are allowable in "repeated applications" of what will be considered to be the same measurement process applied to the measurement of the same quantity under the same conditions. To be realistic, the "allowable variations" must be of sufficient scope to bracket the circumstances likely to be met in practice. Furthermore, any experimental program that aims to determine the standard deviation of a measurement process as an indication of its precision, must be based on appropriate random sampling of this likely range of circumstances.

Ordinarily the accuracy of a measurement process may be characterized by giving (a) the standard deviation of the process and (b) credible bounds to its likely overall systematic error. Determination of credible bounds to the combined effect of recognized potential sources of systematic error always involves some arbitrariness, not only in the placing of reasonable bounds on the systematic error likely to be contributed by each particular assignable cause, but also in the manner in which these individual contributions are combined. Consequently, the "inaccuracy" of end results of measurement cannot be expressed by "confidence limits" corresponding to a definite numerical "confidence level," except in those rare instances in which the possible overall systematic error of a final result is negligible in comparison with its imprecision.

1. Introduction

Calibration of instruments and standards is basically a refined form of measurement. Measurement is the assignment of numbers to material things to represent the relations existing among them with respect to particular properties. One always measures properties of things, not the things themselves. In practice, measurement of some property of a thing ordinarily takes the form of a sequence of steps or operations that yields as an end result a number that indicates how much of this property the thing has, for someone to use for a specific purpose. The end result may be the outcome of a single reading of an instrument. More often it is some kind of average, e.g., the arithmetic mean of a number of independent determinations of the same magnitude, or the final result of a least squares "reduction" of measurements of a number of different quantities that bear known relations to each other in accordance with a definite experimental plan. In general, the purpose for which the answer is needed determines the accuracy required and ordinarily also the method of measurement employed.

Specification of the apparatus and auxiliary equipment to be used, the operations to be performed, the sequence in which they are to be executed, and the conditions under which they are respectively to be carried out—these instructions collectively serve to define a method of measurement. A measurement process is the realization of a method of measurement in terms of particular apparatus and equipment of the prescribed kinds, particular conditions that at best only approximate the conditions prescribed, and particular persons as operators and observers.

It has long been recognized that, in undertaking to apply a particular method of measurement, a degree of consistency among repeated measurements of a single quantity needs to be attained before the method of measurement concerned can be regarded as meaningfully realized, i.e., before a measurement process can be said to have been established that is

---

*Published in J. Research NBS 67C (Eng. and Instr.) No. 2, 161-187 (Apr.-June 1963).

**Chief, Statistical Engineering Section, Applied Mathematics Division, National Bureau of Standards, Washington, D.C.

a realization of the method of measurement concerned. Indeed, consistency or statistical stability of a very special kind is required: to qualify as a measurement process a measurement operation must have attained what is known in industrial quality control language as a state of statistical control. Until a measurement operation has been "debugged" to the extent that it has attained a state of statistical control it cannot be regarded in any logical sense as measuring anything at all. And when it has attained a state of statistical control there may still remain the question of whether it is faithful to the method of measurement of which it is intended to be a realization.

The systematic error, or bias, of a measurement process refers to its tendency to measure something other than what was intended; and is determined by the magnitude of the difference $\mu - \tau$ between the process average or limiting mean μ associated with measurement of a particular quantity by the measurement process concerned and the true value τ of the magnitude of this quantity. On first thought, the "true value" of the magnitude of a particular quantity appears to be a simple straightforward concept. On careful analysis, however, it becomes evident that the "true value" of the magnitude of a quantity is intimately linked to the purposes for which knowledge of the magnitude of this quantity is needed, and cannot, in the final analysis, be meaningfully and usefully defined in isolation from these needs.

The precision of a measurement process refers to, and is determined by the degree of mutual agreement characteristic of independent measurements of a single quantity yielded by repeated applications of the process under specified conditions; and its accuracy refers to, and is determined by, the degree of agreement of such measurements with the true value of the magnitude of the quantity concerned. In brief "accuracy" has to do with closeness to the truth; "precision," only with closeness together.

Systematic error, precision, and accuracy are inherent characteristics of a measurement process and not of a particular measurement yielded by the process. We may also speak of the systematic error, precision, and accuracy of a particular method of measurement that has the capability of statistical control. But these terms are not defined for a measurement operation that is not in a state of statistical control.

The precision, or more correctly, the *im*precision of a measurement process is ordinarily summarized by the standard deviation of the process, which expresses the characteristic disagreement of repeated measurements of a single quantity by the process concerned, and thus serves to indicate by how much a particular measurement is likely to differ from other values that the same measurement process might have provided in this instance, or might yield on remeasurement of the same quantity on another occasion. Unfortunately, there does not exist any single comprehensive measure of the accuracy (or *in*accuracy) of a measurement process analogous to the standard deviation as a measure of its imprecision.

To characterize the accuracy of a measurement process it is necessary, therefore, to indicate (a) its systematic error or bias, (b) its precision (or *im*precision)—and, strictly speaking, also, (c) the form of the distribution of the individual measurements about the process average. Such is the unavoidable situation if one is to concern one's self with individual measurements yielded by any particular measurement process. Fortunately, however, "final results" are ordinarily some kind of average or adjusted value derived from a set of independent measurements, and when four or more independent measurements are involved, such adjusted values tend to be normally distributed to a very good approximation, so that the accuracy of such final results can ordinarily be characterized satisfactorily by indicating (a) their imprecision as expressed by their standard error, and (b) the systematic error of the process by which they were obtained.

The error of any single measurement or adjusted value of a particular quantity is, by definition, the difference between the measurement or adjusted value concerned and the true value of the magnitude of this quantity. The error of any particular measurement or adjusted value is, therefore, a fixed number; and this number will ordinarily be unknown and unknowable, because the true value of the magnitude of the quantity concerned is ordinarily unknown and unknowable. Limits to the error of a single measurement or adjusted value may, however, be inferred from (a) the precision, and (b) bounds on the systematic error of the measurement process by which it was produced—but not without risk of being incorrect, because, quite apart from the inexactness with which bounds are commonly placed on a systematic error of a measurement process, such limits are applicable to the error of the single measurement or adjusted value, not as a unique individual outcome, but only as a typical case of the errors characteristic of such measurements of the same quantity that might have been, or might be, yielded by the same measurement process under the same conditions.

Since the precision of a measurement process is determined by the characteristic "closeness together" of successive independent measurements of a single magnitude generated by repeated application of the process under specified conditions, and its bias or systematic error is determined by the direction and amount by which such measurements tend to differ from the true value of the magnitude of the quantity concerned, it is necessary to be clear on what variations of procedure, apparatus, environmental conditions, observers, etc., are allowable in "repeated applications" or what will be considered to be the same measurement process applied to the measurement of the same quantity under the same conditions. If whatever measures of the precision and bias of a measurement process we may adopt are to provide a realistic indication of the accuracy of this process in practice, then the "allowable variations" must be of sufficient scope to bracket the range of circumstances commonly met in practice. Furthermore, any experimental program that aims to determine the precision, and thence the accuracy of a measurement process, must be based on an appropriate random sampling of this "range of circumstances," if the usual tools of statistical analysis are to be strictly applicable.

When adequate random sampling of the appropriate "range of circumstances" is not feasible, or even possible, then it is necessary (a) to compute, by extrapolation from available data, a more or less subjective estimate of the precision of the measurement process concerned, to serve as a substitute for a direct experimental measure of this characteristic, and (b) to assign more or less subjective bounds to the systematic error of the measurement process. To the extent that such at least partially subjective computations are involved, the resulting evaluation of the overall accuracy of a measurement process "is based on subject-matter knowledge and skill, general information, and intuition—but not on statistical methodology" [Cochran et al. 1953, p. 693]. Consequently, in such cases the statistically precise concept of a family of "confidence intervals" associated with a definite "confidence level" or "confidence coefficient" is not applicable.

The foregoing points and certain other related matters are discussed in greater detail in the succeeding sections, together with an indication of procedures for the realistic evaluation of precision and accuracy of established procedures for the calibration of instruments and standards that minimize as much as possible the subjective elements of such an evaluation. To the extent that complete elimination of the subjective element is not always possible, the responsibility for an important and sometimes the most difficult part of the evaluation is shifted from the shoulders of the statistician to the shoulders of the subject matter "expert."

2. Measurement

2.1. Nature and Object

Measurement is the assignment of numbers to material things to represent the relations existing among them with respect to particular properties. The number assigned to some particular property serves to represent the relative amount of this property associated with the object concerned.

Measurement always pertains to properties of things, not to the things themselves. Thus we cannot measure a meter bar, but can and usually do, measure its length; and we could also measure its mass, its density, and perhaps, also its hardness.

The object of measurement is twofold: first, symbolic representation of properties of things as a basis for conceptual analysis; and second, to effect the representation in a form amenable to the powerful tools of mathematical analysis. The decisive feature is symbolic representation of properties, for which end numerals are not the only usable symbols.

In practice the assignment of a numerical magnitude to a particular property of a thing is ordinarily accomplished by comparison with a set of standards, or by comparison either of the quantity itself, or of some transform of it, with a previously calibrated scale. Thus, length measurements are usually made by directly comparing the length concerned with a calibrated bar or tape; and mass measurements, by directly comparing the weight of a given mass with the weight of a set of standard masses, by means of a balance; but force measurements are usually carried out in terms of some transform, such as by reading on a calibrated scale the extension that the force produces in a spring, or the deflection that it produces in a proving ring; and temperature measurements are usually performed in terms of some transform, such as by reading on a calibrated scale the expansion of a column of mercury, or the electrical resistance of a platinum wire.

2.2. Qualitative and Quantitative Aspects

As Walter A. Shewhart, father of statistical control charts, has remarked:

> "It is important to realize . . . that there are two aspects of an operation of measurement; one is quantitative and the other qualitative. One consists of *numbers* or pointer readings such as the observed lengths in n measurements of the length of a line, and the other consists of the *physical manipulations* of physical things by *someone* in accord with instructions that we shall assume to be describable in words constituting a text." [Shewhart 1939, p. 130.]

More specifically, the qualitative factors involved in the measurement of a quantity are: the *apparatus* and *auxiliary equipment* (e.g., reagents, batteries or other source of electrical energy, etc.) employed; the *operators* and *observers*, if any, involved; the *operations* performed, together with the *sequence* in which, and the *conditions* under which, they are respectively carried out.

2.3. Correction and Adjustment of Observations

The numbers obtained as "readings" on a calibrated scale are ordinarily the end product of everyday measurement in the trades and in the home. In scientific work there are usually two important additional quantitative aspects of measurement: (1) *correction* of the readings, or their transforms, to compensate for known deviations from ideal execution of the prescribed operations, and for non-negligible effects of variations in uncontrolled variables; and (2) *adjustment* of "raw" or corrected measurements of particular quantities to obtain values of these quantities that conform to restrictions upon, or interrelations among, the magnitudes of these quantities imposed by the nature of the problem.

Thus, it may not be practicable or economically feasible to take readings at exactly the prescribed temperatures; but quite practicable and feasible to bring and hold the temperature within narrow neighborhoods of the prescribed values and to record the actual temperatures to which the respective readings correspond. In such cases, if the deviations from the prescribed temperatures are not negligible, "temperature *corrections*" based on appropriate theory are usually applied to the respective readings to bring them to the values that presumably would have been observed if the temperature in each instance had been exactly as prescribed.

In practice, however, the objective just stated is rarely, if ever, actually achieved. Any "temperature corrections" applied could be expected to bring the respective readings "to the values that presumably would have been observed if the temperature in each instance had been exactly as prescribed" if and only if these "temperature corrections" made appropriate allowances for *all* of the effects of the deviations of the actual temperatures from those prescribed. "Temperature corrections" ordinarily correct *only* for particular effects of the deviations of the actual temperatures from their prescribed values; *not* for all of the effects on the readings traceable to deviations of the actual temperatures from those prescribed. Thus Michelson utilized "*temperature* corrections" in his 1879 investigation of the speed of light; but his results exhibit a dependence on temperature after "temperature correction." The "temperature corrections" applied corrected only for the effects of thermal expansion due to variations in temperature and not also for changes in the index of refraction of the air due to changes in the humidity of the air, which in June and July at Annapolis is highly correlated with temperature. *Corrections applied in practice are usually of more limited scope than the names that they are given appear to indicate.*

Adjustment of observations is fundamentally different from their "correction." When two or more related quantities are measured individually, the resulting measured values usually fail to satisfy the constraints on their magnitudes implied by the given interrelations among the quantities concerned. In such cases these "raw" measured values are mutually contradictory, and require *adjustment* in order to be usable for the purpose intended. Thus, measured values of the three cyclic differences $(A-B)$, $(B-C)$, and $(C-A)$ between the lengths of three nominally equivalent gage blocks are mutually contradictory, and strictly speaking are not usable as values of these differences, unless they sum to zero.

The primary goal of *adjustment* is to derive from such inconsistent measurements, if possible, *adjusted values* for the quantities concerned that do satisfy the constraints on their magnitudes imposed by the nature of the quantities themselves and by the existing interrelations among them. A second objective is to select from all possible sets of adjusted values the set that is the "best"—or, at least, a set that is "good enough" for the intended purpose—in some well-defined sense. Thus, in the above case of the measured differences between the lengths of three gage blocks, an adjustment could be effected by ignoring the measured value of one of the differences entirely, say, the difference $(C-A)$, and taking the negative of the sum of the other two as its adjusted value,

$$Adj(C-A)=-[(A-B)+(B-C)].$$

This will certainly assure that the sum of all three values, $(A-B)+(B-C)+Adj(C-A)$, is zero, as required, and is clearly equivalent to ascribing all of the excess or deficit to the replaced measurement, $(C-A)$. Alternatively, one might prefer to distribute the necessary total adjustment $-[(A-B)+(B-C)+(C-A)]$ equally over the individual measured differences, to obtain the following set of adjusted values:

$$Adj\,(A-B)=(A-B)-\frac{1}{3}[(A-B)+(B-C)+(C-A)]$$

$$=\frac{1}{3}[2(A-B)-(B-C)-(C-A)]$$

$$Adj\,(B-C)=\frac{1}{3}[2(B-C)-(A-B)-(C-A)]$$

$$Adj\,(C-A)=\frac{1}{3}[2(C-A)-(A-B)-(B-C)].$$

Clearly, the sum of these three adjusted values must always be zero, as required, regardless of the values of the original individual measured differences. Furthermore, most persons, I believe, would consider this latter adjustment the better; and under certain conditions with respect to the "law of error" governing the original measured differences, it is indeed the "best."

Note that no adjustment problem existed at the stage when only two of these differences had been measured whichever they were, for then the third could be obtained by subtraction. As a general principle, when no more observations are taken than are sufficient to provide one value of each of the unknown quantities involved, then the results so obtained are usable at least—they may not be "best." On the other hand, when additional observations are taken, leading to "over determination" and consequent contradiction of the fundamental properties of, or the basic relationships among the quantities concerned, then the respective observations must be regarded as contradicting one another. When this happens the observations themselves, or values derived from them, must be replaced by adjusted values such that all contradiction is removed. "This is a logical necessity, since we cannot accept for truth that which is contradictory or leads to contradictory results." [Chauvenet 1868, p. 472.]

2.4. Scheduling the Taking of Measurements

Having done what one can to remove extraneous sources of error, and to make the basic measurements as precise and as free from systematic error as possible, it is frequently possible not only to increase the precision of the end results of major interest but also to simultaneously decrease their sensitivity to sources of possible systematic error, by careful scheduling of the measurements required. An instance is provided by the traditional procedure for calibrating liquid-in-glass thermometers [Waidner and Dickinson 1907, p. 702; NPL 1957, pp. 29–30; Swindells 1959, pp. 11–12]: Instead of attempting to hold the temperature of the comparison bath constant, a very difficult objective to achieve, the heat input to the bath is so adjusted that its temperature is slowly increasing at a steady rate, and then readings of, say, four test thermometers and two standards are taken in accordance with the schedule

$$S_1T_1T_2T_3T_4S_2S_2T_4T_3T_2T_1S_1$$

the readings being spaced uniformly in time so that the arithmetic mean of the two readings of any one thermometer will correspond to the temperature of the comparison bath at the midpoint of the period. Such scheduling of measurement taking operations so that the effects of the specific types of departures from perfect control of conditions and procedure will have an opportunity to balance out is one of the principal aims of the art and science of *statistical design of experiments*. For additional physical science examples, see, for instance, Youden [1951a; and 1954–1959].

2.5. Measurement as a Production Process

We may summarize our discussion of measurement up to this point, as follows: Measurement of some property of a thing in practice always takes the form of a sequence of steps or operations that yield as an end result a number that serves to represent the amount or quantity of some particular property of a thing—a number that indicates how much of this property the thing has, for someone to use for a specific purpose. The end result may be the outcome of a single reading of an instrument, with or without corrections for departures from prescribed conditions. More often it is some kind of average or adjusted value, e.g., the arithmetic mean of a number of independent determinations of the same magnitude, or the final result of, say, a least squares "reduction" of measurements of a number of different quantities that have known relations to the quantity of interest.

Measurement of some property of a thing is ordinarily a repeatable operation. This is certainly the case for the types of measurement ordinarily met in the calibration of standards and instruments. It is instructive, therefore, to regard measurement as a *production process*, the "product" being the numbers, that is. the measurements that it yields; and to compare and contrast measurement processes in the laboratory with mass production processes in industry. For the moment it will suffice to note (a) that when successive amounts of units of "raw material" are processed by a particular mass production process, the output is a series of nominally identical items of product—of the particular type produced by the mass production operation, i.e., by the *method of production* concerned; and (b) that when successive objects are measured by a particular measurement process, the individual items of "product" produced consist of the numbers assigned to the respective objects to represent the relative amounts that they possess of the property determined by the *method of measurement* involved.

2.6. Methods of Measurement and Measurement Processes

Specification of the apparatus and auxiliary equipment to be used, the operations to be performed, the sequence in which they are to be carried out, and the conditions under which they are respectively to be carried out—these *instructions* collectively serve to define a *method of measurement*. To the extent that corrections may be required they are an integral part of measurement. The types of corrections that will ordinarily need to be made, and specific procedures for making them, should be included among "the operations to be performed." Likewise, the essential adjustments required should be noted, and specific procedures for making them incorporated in the specification of a method of measurement.

A *measurement process* is the realization of a method of measurement in terms of particular apparatus and equipment of the prescribed kinds, particular conditions that at best only approximate the conditions prescribed, and particular persons as operators and observers [ASTM 1961, p. 1758; Murphy 1961, p. 264]. Of course, there will often be a question whether a particular measurement process is loyal to the method of measurement of which it is intended to be a realization; or whether two different measurement processes can be considered to be realizations of the same method of measurement.

To begin with, written specifications of methods of measurement often contain absolutely precise instructions which, however, cannot be carried out (repeatedly) with complete exactitude in practice; for example, "move the two parallel cross hairs of the micrometer of the microscope until the graduation line of the standard is centered between them." The accuracy with which such instructions can be carried out in practice will always depend upon "the circumstances"; in the case cited, on the skill of the operator, the quality of the graduation line of the standard, the quality of the screw of the micrometer, the parallelism of the cross hairs, etc. To the extent that the written specification of a method of measurement involves absolutely precise instructions that cannot be carried out with complete exactitude in practice there are certain to be discrepancies between a method of measurement and its realization by a particular measurement process.

In addition, the specification of a method of measurement often includes a number of imprecise instructions, such as "raise the temperature slowly," "stir well before taking a reading," "make sure that the tubing is clean," etc. Not only are such instructions inherently vague, but also in any given instance they must be understood in terms of the general level of refinement characteristic of the context in which they occur. Thus, "make sure that the tubing is clean" is not an absolutely definite instruction; to some people this would mean simply that the tubing should be clean enough to drink liquids through; in some laboratory work it might be interpreted to mean mechanically washed and scoured so as to be free from dirt and other ordinary solid matter (but not cleansed also with chemical solvents to remove more stubborn contaminants); to an advanced experimental physicist it may mean not merely mechanically washed and chemically cleansed, but also "out gassed" by being heated to and held at a high temperature, near the softening point, for an hour or so. All will agree, I believe, that it would be exceedingly difficult to make such instructions absolutely definite with a convenient number of words. To the extent that the specification of a method of measurement includes instructions that are not absolutely definite, there will be room for differences between measurement processes that are intended to be realization of the very same method of measurement.

Recognition of the difficulty of achieving absolute definiteness in the specification of a method of measurement does not imply that "any old set" of instructions will serve to define a method of measurement. Quite the contrary. To qualify as a specification of a method of measurement, a set of instructions must be sufficiently definite to insure statistical stability of repeated measurements of a single quantity, that is, derived measurement processes must be capable of meeting the criteria of *statistical control* [Shewhart 1939, p. 131; Murphy 1961, p. 265; ASTM 1961, p. 1758]. To elucidation of the meaning of, and need for this requirement we now turn.

3. Properties of Measurement Processes

3.1. Requirement of Statistical Control

The need for attaining a degree of consistency among repeated measurements of a single quantity before the method of measurement concerned can be regarded as meaningful has certainly been recognized for a long, long time. Thus Galileo, describing his famous experiment on the acceleration of gravity in which he allowed a ball to roll different distances down an inclined plane wrote:

> ". . . si lasciava (como dico) scendere per il detto canale la palla, notando, nel modo che appresso dirò, il temp che consumava nello scorrerlo tutto, replicando il medesimo atto molte volte per assicurarsi bene della quantità del temp, nel quale non si trovava mai differenza nè anco della decima parte d'una battuta di polso. Fatta e stabilita precisamente tale operazione, facemmo scender la medisima palla solamente per la quarta parte della lunghezza di esso canale . . ."[1] [Galileo 1638, Third Day; Nat'l. ed., p. 213.]

Something more than mere "consistency" is required, however, as Shewhart points out eloquently in his very important chapter on "The Specification of Accuracy and Precision" [Shewhart 1939, ch. IV]. He begins by noting that the description given by R. A. Millikan [1903, pp. 195–196] of a method for determining the surface tension T of a liquid from measurements of the force of tension F of a film of the liquid contains the following instruction with regard to the basic readings from which measurements of F are derived: "Continue this operation until a number of consistent readings can be obtained." Shewhart then comments on this as follows:

> ". . . the text describing the operation does not say to carry out such and such physical operations and call the result a measurement of T. Instead, it says in effect not to call the result a measurement of T until one has attained a certain degree of *consistency* among the observed values of F and hence among those of T. Although this requirement is not always explicitly stated in specifications of the operation of measurements as it was here, I think it is always implied. Likewise, I think it is always assumed that there can be too much consistency or uniformity among the observed values as, for example, it a large number of measurements of the surface tension of a liquid were found to be identical. What is wanted but not explicitly described is a specific kind and degree of consistency.
>
> ". . . it should be noted that the advice to repeat the operation of measuring surface tension until a number of consistent readings have been obtained is indefinite in that it does not indicate how many readings shall be taken before applying a test for consistency, nor what kind of test of consistency is to be applied to the numbers or pointer readings . . . . One of the objects of this chapter is to see how far one can go toward improving this situation by providing an operationally definite criterion that preliminary observations must meet before they are to be considered consistent in the sense implied in the instruction cited above.
>
> "Before doing this, however, we must give attention not so much to the consistency of the n observed values already obtained by n repetitions of the operation of measurement as we do to the *reproducibility of the operation* as determined by the numbers in the potentially infinite sequence corresponding to an infinite number of repetitions of this operation. No one would care very much how consistent the first n preliminary observations were if nothing could be validly inferred from this as to what future observations would show. Hence, it seems to me that the characteristics of the numerical aspects of an operation that is of greatest practical interest is its *reproducibility within tolerance limits throughout the infinite sequence*. The limit to which we may go in this direction is to attain a state of statistical control. The attempt to attain a certain kind of consistency within the first n observed values is merely a means of attaining reproducibility within limits throughout the whole of the sequence." [Shewhart 1939, pp. 131–132.]

The point that Shewhart makes forcefully, and stresses repeatedly later in the same chapter, is that the first n measurements of a given quantity generated by a particular measurement process provide a logical basis for predicting the behavior of further measurements of the same quantity by the same measurement process if and only if these n measurements may be regarded as a *random sample* from a "population" or "universe" of all conceivable measurements of the given quantity by the measurement process concerned; that is, in the language of mathematical statistics, if an only if the n measurements in hand may be regarded as "observed values" of a sequence of random variables characterized by a probability distribution identified with the measurement process concerned, and related through the values of one or more of its parameters to the magnitude of the quantity measured.

It should be noted especially that nothing is said about the mathematical form of the *probability distribution* of these random variables. The important thing is that there be one. W. Edwards

---

[1] I am grateful to my colleague Ugo Fano for the following literal translation: ". . . we let, as I was saying, the ball descend through said channel, recording, in a manner presently to be described, the time it took in traversing it all, repeating the same action many times to make really sure of the magnitude of time, in which one never found a difference of even a tenth of a pulsebeat. Having done and established precisely such operation, we let the same ball descend only for the fourth part of the length of the same channel; . . ."

Deming has put this clearly and forcefully in these words:

> "In applying statistical theory, the main consideration is not what the shape of the universe is, but whether there is any universe at all. No universe can be assumed, nor . . . statistical theory . . . applied unless the observations show statistical control. In this state the samples when cumulated over a suitable interval of time give a distribution of a particular shape, and this shape is reproduced hour after hour, day after day, so long as the process remains in statistical control—i.e., exhibits the properties of randomness. In a state of control, *n* observations may be regarded as a sample from the universe of whatever shape it is. A big enough sample, or enough small samples, enables the statistician to make meaningful and useful predictions about future samples. This is as much as statistical theory can do.
>
> ". . . Very often the experimenter, instead of rushing in to apply [statistical methods] should be more concerned about attaining statistical control and asking himself whether any predictions at all (the only purpose of his experiment), by statistical theory or otherwise, can be made." [Deming 1950, pp. 502–503.]

Shewhart was well aware of the fact that from a set of *n* measurements in hand it is not possible to decide with absolute certainty whether they do or do not constitute a *random sample* from some definite statistical "population" characterized by a probability distribution. He, therefore, proposed [Shewhart 1939, pp. 146–147] that in any particular instance one should "decide to act for the present as if"[2] the measurements in hand (and their immediate successors) were a simple random sample from a definite statistical population—i.e., in the language of mathematical statistics, were "observed values" of *independent identically distributed random variables*—only if the measurements in hand met the requirements of the small-samples version of Criterion I of his previous book [Shewhart 1931, pp. 309–318] and of certain additional tests of randomness that he described explicitly for the first time in his contribution to the University of Pennsylvania Bicentennial Conference in September 1940 [Shewhart, 1941]. In other words, Shewhart proposed that one should consider a measurement process to be—i.e., should "decide to act for the present as if" the process were—in a *state of (simple) statistical control*, only if the measurements in hand show no evidence of lack of statistical control when analyzed for randomness *in the order in which they were taken* by the control chart techniques for averages and standard deviations that he had found so valuable in industrial process control and by certain additional tests for randomness based on "runs above and below average" and "runs up and down."[3]

Simpson[3] did not prove that taking of the Arithmetic Mean was the best thing to do but merely that it is good. However, in accomplishing this goal he did something much more important: he took the bold step of regarding errors of measurement, not as unique unrelated magnitudes unamenable to mathematical analysis, but as distributed in accordance with a probability distribution that was an intrinsic property of the measurement process itself. He thus opened the way to a mathematical theory of measurement based on the mathematical theory of probability; and, in particular, to the formulation and development of the Method of Least Squares in essentially its present day form by Gauss (1809, 1821) and Laplace (1812).

"Student" (William Sealy Gosset, 1876–1937), pioneer statistical consultant and "father" of the "theory of small samples," was certainly among the first to stress the importance of randomness in measurement and experimentation. Thus, he began his revolutionary 1908 paper on "The probable error of a mean" with these remarks:

> "Any experiment may be regarded as forming an individual of a 'population' of experiments which might be performed under the same conditions. A series of experiments is a sample drawn from this population.
>
> "Now any series of experiments is only of value in so far as it enables us to form a judgment as to the statistical constants of the population to which the experiments belong." [Student 1908, p. 1.]

None of these writers, nor any of their contemporaries, however, provided "an operationally definite criterion that preliminary observations must meet" before we take it upon ourselves "to act for the present as if" they and their immediate successors were random samples from a "population" or "universe" of all conceivable measurements of the given quantity by the measurement process concerned. Provision of such a criterion is Shewhart's major contribution.

Experience shows that in the case of measurement processes the ideal of strict statistical control that Shewhart prescribes is usually very difficult to attain, just as in the case of industrial production processes. Indeed, many measurement processes simply do not and, it would seem, cannot be made to conform to this ideal of producing successive measurements of a single quantity that can be considered to be "observed values" of independent identically distributed random variables.[4] The nature of the "trouble" was stated succinctly by Student in 1917 when, speaking of physical and chemical determinations, he wrote:

> "After considerable experience I have not encountered any determination which is not influenced by the date on which it is made; from this it follows that a number of determinations of the same thing made on the same day are likely

---

[2] This very explicit phraseology is due to John W. Tukey [1960, p. 424].

[3] Thomas Simpson, in his now famous letter [Simpson 1755] to the President of the Royal Society of London "on the Advantage of taking the Mean of a Number of Observations, in practical Astronomy," was the first to consider repeated measurements of a single quantity by a given measurement process as observed values of independent random variables having the same probability distribution. His conclusion is of interest in itself:

> "Upon the whole of which it appears, that the taking of the Mean of a number of observations, greatly diminishes the chances for all the smaller errors, and cuts off almost all possibility of any great ones: which last consideration, alone, seem sufficient to recommend the use of the method, not only to astronomers, but to all others concerned in making of experiments of any kind (to which the above reasoning is equally applicable). And the more observations or experiments there are made, the less will the conclusion be liable to err, provided they admit of being repeated under the same circumstances."

[4] Looking at the matter from a fundamental viewpoint, perhaps we should say, not that Shewhart's ideal of strict statistical control is unattainable in the case of such measurement processes, but rather that the degree of approximation to this ideal can be made as close as one chooses, if one is willing to pay the price. In other words, how close one chooses to bring a measurement process to the ideal of strict statistical control is, in any given instance, basically an economic matter, taking into account, of course, not only the immediate purpose(s) for which the measurements are intended but also the other uses to which they may be put. (Compare Simon [1946, p. 566] and Eisenhart [1952, p. 554)].

> to lie more closely together than if the repetitions had been made on different days." [Student 1917, p. 415.]

In other words, production of measurements seems to be like the production of paint; and just as in the case of paint, if one must cover a large surface all of which is visible simultaneously, one will do well to use paint all from the same batch, so in the case of measurements, if a scientist or metrologist "wishes to impress his clients" he will "arrange to do repetition analyses as nearly as possible at the same time." [Student 1927, p. 155.]

Fortunately, just as one may blend paint from several batches to obtain a more uniform color, and one which is, presumably, closer to the "process average," so also may a scientist or metrologist "if he wishes to diminish his real error, . . . separate [his measurements] by as wide an interval of time as possible" [Student, loc. cit.] and then take an appropriate average of them as his determination. Consequently, if we are to permit such averaging as an allowable step in a fully specified measurement process (see sec. 2.6 above), then we are obliged to recognize both within-day and between-day components of variation, and accept such a complex measurement process as being in a state of statistical control overall, or as we shall say, in a *state of COMPLEX statistical control,* when the components of within-day and between-day variation are both in a state of statistical control in Shewhart's strict sense, which we shall term *SIMPLE statistical control.* In more complex situations, one may be obliged to recognize more than two "layers" of variation, and, sometimes, more than a single component of variation within a given "layer."

Adopting this more general concept of statistical control, R. B. Murphy of the Bell Telephone Laboratories in his essay "On the Meaning of Precision and Accuracy" [Murphy 1961], published in advance of the issuance by the American Society for Testing and Materials of its Tentative Recommended Practice with respect to the "Use of the Terms Precision and Accuracy as Applied to Measurement of a Property of a Material" [ASTM 1961], remarks:

> "Following through with this line of thought borrowed from quality control, we shall add a requirement that an effort to follow a test method ought not to be known as a measurement process unless it is capable of statistical control. Capability of control means that either the measurements are the product of an identifiable statistical universe or an orderly array of such universes or, if not, the physical causes preventing such identification may themselves be identified and, if desired, isolated and suppressed. Incapability of control implies that the results of measurement are not to be trusted as indications of the physical property at hand—in short, we are not in any verifiable sense measuring anything . . . . Without this limitation on the notion of measurement process, one is unable to go on to give meaning to those statistical measures which are basic to any discussion of precision and accuracy." [Murphy 1961, pp. 264–265.]

3.2. Postulate of Measurement and the Concept of a Limiting Mean

A conspicuous characteristic of measurement is disagreement of repeated measurements of the same quantity. Experience shows that, when high accuracy is sought, repeated measurements of the same quantity by a particular measurement process does not yield uniformly the same number.[5] We explain these discordances by saying that the individual measurements are affected by *errors,* which we interpret to be the manifestations of variations in the execution of the process of measurement resulting from "the imperfections of instruments, and of organs of sense," and from the difficulty of achieving (or even specifying with a convenient number of words) the ideal of perfect control of conditions and procedure.

This "cussedness of measurements" brings us face to face with a fundamental question: In what sense can we say that the measurements yielded by a particular measurement process serve to determine a unique magnitude, when experience shows that repeated measurement of a single quantity by this process yields a sequence of nonidentical numbers. What is the value thus determined?

The answer takes the form of a *postulate* about measurement processes that has been expressed by N. Ernest Dorsey, as follows:

> "The mean of a family of measurements—of a number of measurements for a given quantity carried out by the same apparatus, procedure and observer—approaches a definite value as the number of measurements is indefinitely increased. Otherwise, they could not properly be called measurements of a given quantity. In the theory of errors, this limiting mean is frequently called the 'true' value, although it bears no necessary relation to the true quaesitum, to the actual value of the quantity that the observer desires to measure. This has often confused the unwary. Let us call it the limiting mean." [Dorsey 1944, p. 4; Dorsey and Eisenhart 1953, p. 103.]

In my lectures at the National Bureau of Standards, and elsewhere, I have termed this—or rather a slightly rephrased version of it—the *Postulate of Measurement.* A mathematical basis for it is provided by the Strong Law of Large Numbers, a theorem in the mathematical theory of probability discovered during the present century. See, for example, Feller [1957, pp. 243–245, 374], Gnedenko [1962, pp. 241–249], or Parzen [1960, p. 420].

Needless to say, by a "family of measurements" Dorsey means, not a succession of "raw" readings, but rather a succession of adjusted or corrected values which, by virtue of adjustment or correction, can rightfully be considered to be determinations of a single magnitude.

a. Mathematical Formulation

The foregoing can be expressed mathematically as follows: on some particular occasion, say the *i*th, we may take a number of successive measurements of a single quantity by a given measurement process under certain specified circumstances. Let

$$x_{i1},\ x_{i2},\ .\ .\ .\ ,\ x_{ij},\ .\ .\ . \qquad (1)$$

---

[5] The qualification "when high accuracy is sought" is essential; for if using an ordinary two-pan chemical balance we measure and record the mass of a small metallic object only to the nearest gram, then we would expect all of our measurements to be the same—except in the equivocal case of a mass equal, or very nearly equal, to an odd multiple of ½ g, and such equivocal cases can be resolved easily by adding a ½ g mass to one pan. Full accordance of measurements clearly cannot be taken as incontestable evidence of high accuracy; but rather should be regarded as evidence of limited accuracy.

denote the sequence of measurements so generated. Conceptually at least, this sequence could be continued indefinitely. Likewise, on different occasions we might start a new sequence, using the same measurement procedure and applying it to measurement of the same quantity under the same fixed set of circumstances. Each such fresh "start" would correspond to a different value of i. If, for example, the measurement process concerned is statistically stable in the sense of being in a *state of statistical control* as defined by Shewhart [1939], then the Strong Law of Large Numbers will be applicable and we may expect the sequence of cumulative arithmetic means on the ith occasion, namely,

$$\bar{x}_{in} \equiv (x_{i1} + x_{i2} + \ldots + x_{in})/n, \ (n=1, 2, \ldots), \quad (2)$$

to converge to μ, a number that constitutes the limiting mean associated with the quantity measured by this measurement process under the circumstances concerned, but independent of the "occasion," that is, independent of the value of "i." The Strong Law of Large Numbers does not guarantee that the sequence (2) for a particular value of "i" will converge to μ as the number of observations n on this occasion tends to infinity, but simply states that among the family of such sequences corresponding to a large number of different starts, $(i=1, 2, \ldots)$, the *instances of nonconvergence to* μ will be *rare exceptions*. In other words, if the measurement process with which one is concerned satisfies the conditions for validity of the Strong Law of Large Numbers, then in practice one is almost certain to be working with a "good" sequence—one for which (2) would converge to μ if the number of observations were continued indefinitely—but "bad" occasions can occur, though rarely. Thus, the Postulate of Measurement expresses something better than an "on-the-average" property—it expresses an "in-almost-all-cases" property. Furthermore, this limiting mean μ, the value of which each individual measurement x is trying to express, can be regarded not only as the *mean* or "center of gravity" of the infinite conceptual population of all measurements x that might conceivably be generated by the measurement process concerned under the specified circumstances, but also as *the* value of the quantity concerned as determined by *this* measurement process.

b. Aim of the Postulate

The sole aim of the Postulate of Measurement is axiomatic acceptance of the existence of a limit approached by the arithmetic mean of a finite number n of measurements generated by any measurement process as $n \rightarrow \infty$. It says nothing about how the "best" estimate of this limiting mean is to be obtained from a finite number of such observations. The Postulate is an answer to the need of the practical man for a justification of his desire to consider the sequence of nonidentical numbers that he obtains when he attempts to measure a quantity "by the same method under like circumstances" as pertaining to a single magnitude, in spite of the evident discordance of its elements. The Postulate aims to satisfy this need by telling him that if he were to continue taking more and still more measurements on this quantity "by the same method under like circumstances" ad infinitum, and were to calculate their cumulative arithmetic means at successive stages of this undertaking, then he would find that the successive terms of this sequence of cumulative arithmetic means would settle down to a narrower and ever narrower neighborhood of some definite number which he could then accept as *the* value of the magnitude that his first few measurements were striving to express.

c. Importance of Limiting Mean

The concept of a *limiting mean* associated with the measurement of a given quantity by a particular measurement process that is in a *state of statistical control* is important because by means of statistical methods based on the mathematical theory of probability we can make quantitative inferential statements, with known chances of error, about the magnitude of this limiting mean from a set of measurements of the given quantity by the measurement process concerned. The magnitude of the limiting mean associated with the measurement of a given quantity by a particular measurement process must be carefully distinguished from the *true magnitude* of the quantity measured, about which we may be tempted to make similar inferential statements. Insofar as we make statistical inferences from a set of measurements, we make them with respect to a property of the measurement process involved under the circumstances concerned. The step from quantitative inferential statements about the limiting mean associated with the measurement of a given quantity by a particular measurement process, to quantitative statements about the true magnitude of the quantity concerned, may be based on subject matter knowledge and skill, general information and intuition—but not on statistical methodology. (Compare Cochran, Mosteller, and Tukey [1953, pp. 692–693].)

3.3. Definition of the Error of a Measurement, and of the Systematic Error, Precision, and Accuracy of a Measurement Process

a. Error of a Single Measurement or Adjusted Value

The *error* of any measurement of a particular quantity is, by definition, the difference between the measurement concerned and the *true value* of the magnitude of this quantity, taken positive or negative accordingly as the measurement is greater or less than the true value. In other words, if x denotes a single measurement of a quantity, or an adjusted value derived from a specific set of individual measurements, and τ is the *true value* of the magnitude of the quantity concerned, then, by definition,

$$\text{the } \textit{error} \text{ of } x \text{ as a measurement of } \tau \equiv x - \tau.$$

The error of any particular measurement or adjusted value, x, is, therefore, a fixed number. The numerical magnitude and sign of this number will ordinarily be unknown and unknowable, because the true value of the magnitude of the quantity concerned is ordinarily unknown and unknowable. Limits to the error of a single measurement or adjusted value may, however, be inferred from (a) the *precision*, and (b) bounds on the *systematic error*, of the measurement process by which it was produced—but not without risk of being incorrect, because, quite apart from the inexactness with which bounds are commonly placed on the systematic error of a measurement process, such limits are applicable to the error of a single measurement or adjusted value, not as a unique individual outcome, but only as a typical case of the errors characteristic of measurements of the same quantity that might have been, or might be, yielded by the same measurement process under the same conditions.

b. Systematic Error of a Measurement Process

When the limiting mean μ associated with measurement of the magnitude of a quantity by a particular measurement process does not agree with the *true value* τ of the magnitude concerned, the measurement process is said to have a *systematic error*, or *bias*, of magnitude $\mu-\tau$.

The systematic error of a measurement process will ordinarily have both constant and variable components. Consider, for example, measurement of the distance between two points by means of a graduated metal tape [Holman 1892, p. 9]. Possible causes of systematic error that immediately come to mind are:

(1) Mistakes in numbering the scale divisions of the tape;
(2) irregular spacing of the divisions of the tape;
(3) sag of tape;
(4) stretch of tape;
(5) temperature not that for which the tape was calibrated.

For *any single distance*, the effects of (1) and (2) will be constant; and the effects of (3) and (4) will undoubtedly each contain a constant component characteristic of the distance concerned. Some of these effects will be of one sign, some of the other, and their algebraic sum will determine the *constant error* of this measurement process with respect to the particular distance concerned. Furthermore, the "constant error" of this measurement process will be different (at least, conceptually) for different distances measured.

In the case of repeated measurement of a single distance, the effect of (5), and at least portions of the effects of (3) and (4), may be expected to vary from one "occasion" to the next (e.g., from day to day), thus contributing *variable components* to the *systematic error* of the process.

A large fraction of the variable contributions of (3) and (4) could, and in practice no doubt would, be removed by stretching the tape by a spring balance or other means so that it is always under the same tension. The stretch corresponding to a particular distance would then be nearly the same at all times, and a fixed correction could be made for most of the sag corresponding to this distance. Furthermore, the effect of (5) could, and in practice probably would, be reduced by determining the temperature of the tape at various points along its length and applying a temperature correction. By comparison of the tape with a standard, the error arising from (1) could be eliminated entirely, and corrections determined as a basis for eliminating, or at least, reducing the effect of (2).

As in the foregoing example there are usually certain obvious sources of systematic error. Unfortunately, there are generally additional sources of systematic error, the detection, diagnosis, and eradication of which call for much patience and acumen on the part of the observer. The work involved in their detection, diagnosis, and eradication often far exceeds that of taking the final measurements, and is sometimes discouraging to the experienced observer as well as to the beginner. Fortunately, there are various statistical tools that are helpful in this connection, and Olmstead [1952] has found that of these the two most effective and universally useful are the average ($\bar{x}$) and range (R) charts of industrial quality control. (For details on the construction and use of $\bar{x}$- and R-charts, see, for example, the ASTM Manual on Quality Control of Materials [ASTM 1951, pp. 61–63 and p. 83]; or American Standards Z1.2–1958 and Z1.3–1958 [ASA 1958b, ASA 1958c].)

c. Concept of True Value

In the foregoing we have defined the *error* of a measurement x to be the difference $x-\tau$ between the measurement and the *true value* τ of the magnitude of the quantity concerned; and the *systematic error*, or *bias*, of a measurement process as the difference $\mu-\tau$ between the limiting mean μ associated with the measurement of a particular quantity by the measurement process concerned, and the *true value* τ of the magnitude of this quantity. This immediately raises the question: Just how is the "true value" of the magnitude of a particular property of some thing defined? In the final analysis, the "true value" of the magnitude of a quantity is defined by agreement among experts on an *exemplar method* for the measurement of its magnitude—it is the limiting mean of a conceptual *exemplar process* that is an ideal realization of the agreed-upon exemplar method. And the refinement to which one should go in specifying the exemplar process will depend on the purposes for which a determination of the magnitude of the quantity concerned is needed—not just the immediate purpose for which measurements are to be taken but also the other uses to which these measurements, or a final adjusted value derived therefrom, may possibly be put.

Consider, for example, the "true value" of the length of a particular gage block. In our minds we envisage the gage block as a rectangular parallelepiped, and its *length* is, of course, the distance between its two "end" faces. But it is practically certain that the particular gage block in question is not an exact rectangular parallelepiped; and that its two end faces are not planes, nor even absolutely smooth surfaces. Shall we define the "true length" of this gage block to be the distance between the "tops" of the highest "mountains" at each end, i.e., the distance between the two "outermost points" at each end? If so, is this distance to be measured diagonally, if necessary, or parallel to the "lengthwise axis" of the gage block? If the latter, then we have the problem of how this "length-wise axis" is to be defined, especially in the case of a thin gage block whose *length* corresponds to what would ordinarily be considered to be its thickness. Or shall we be, perhaps, more sophisticated, and envisage a "mean plane" at each end, which in general will not be parallel to each other, and define the length of this gage block to be the distance between two particular points on these planes. If we choose the "outermost points" we again have the problem of the direction in which the distance is to be measured. Alternatively, we might define the length of this gage block to be the distance between two strictly parallel and conceptually perfect optical flats "just touching" the gage block at each end. If so, then is the "true distance" between these flats defined in terms of wavelengths of light via the techniques of optical interferometry the "true length" of the gage block appropriate to the purposes for which the gage block is to be used, namely, to calibrate gages and to determine the lengths of other objects by *mechanical* comparisons? Furthermore, it is clear, that the intrinsic difficulty of defining the "true value" of the *length* of a particular gage block is not eliminated if, instead, we undertake to define the "true value" of the *difference in length* of two particular gage blocks, one of which is a standard, the *accepted value* of whose length is, say, *m* microinches *exactly*, by industry, national or international agreement.

Similar difficulties arise, of course, in the definition of the "true value" of the *mass* of a mass standard, one of which has been resolved by international agreement. In defining the "true value" of the *mass* of a particular metallic mass standard, shall *the* mass of this particular standard be envisaged as the mass of its metallic substance alone, relative to the International Prototype Kilogram, or as the mass of its metallic substance plus the mass of the air and water vapor adsorbed upon its surface under standard conditions? The difference amounts to about 45 μg in the case of a platinum-iridium standard kilogram, and becomes critical in the case of 500 mg standards. *The mass* of a mass standard is, therefore, specified in measurement science to be the mass of the metallic substance of the standard *plus* the mass of the average volume of air adsorbed upon its surface under standard conditions. Definition of the "true value" of the *mass* of a mass standard, and *a fortiori*, of the *difference in mass* of two mass standards is, therefore, a very complex matter.

W. Edwards Deming uses the expression "preferred procedure" for what we have termed an "exemplar method," and very sagely remarks that "a preferred procedure is distinguished by the fact that it supposedly gives or would give results nearest to what are needed for a particular end; and also by the fact that it is more expensive or more time consuming, or even impossible to carry out," adding that "as a preferred procedure is always subject to modification or obsolescence, we are forced to conclude that *neither the accuracy nor the bias of any procedure can ever be known in a logical sense.*" [Deming 1950, pp. 15–17.]

It should be evident from the foregoing that the "true value" of the magnitude of some property of a thing or system cannot be defined with complete absolute exactitude.

As Cassius J. Keyser has remarked, "Absolute certainty is a privilege of uneducated minds—and fanatics. It is, for scientific folk, an unattainable ideal." [Keyser 1922, p. 120.] The degree of refinement to which one will, or ought, to go in a particular instance will depend on the uses for which knowledge of the magnitude of the property concerned is needed. The "true value" of the length of a piece of cloth in everyday commerce is certainly a fuzzy concept. "Certainly we are not going to specify that the cloth shall be measured while suspended horizontally under a tension of *x* pounds, at an ambient temperature of *y* degrees and a relative humidity of *z* percent" [Simon 1946, p. 654]. On the other hand, a moderate degree of refinement is necessary in defining the "true length" and "true width" of the recessed area in a window sash to which a pane of glass is to be fitted. Considerably greater refinement is needed in the definition of the "true value" of the *length* of a gage block, of the *mass* of a mass standard or of the *frequency* of a frequency standard—and in the last mentioned case there is not today, I understand, complete agreement among experts on the matter.

Indeed, as is evident from the foregoing, the "true value" of the magnitude of a particular quantity is intimately linked to the purposes for which a value of the magnitude of this quantity is needed, and its "true value" cannot, in the final analysis, be defined meaningfully and usefully in isolation from these needs. Therefore, as this fact becomes more widely recognized in science and engineering, I hope that the traditional term "true value" will be discarded in measurement theory and practice, and replaced by some more appropriate term such as "target value"[6] that conveys the idea of being the value that one would like to obtain for the purpose in hand, without any implication that it is some sort of permanent constant preexisting and transcending any use that we may have for it. I have retained the traditional expression "true value" in the sequel because of its greater familiarity, but shall always mean by it the relevant "target value."

---

[6] "We admit the existence of systematic error—of a difference between the quantity measured (the measured quantity) and the quantity of interest (the target quantity). We ask the observations about the measured quantity. We ask our subject matter knowledge, intuition, and general information about the relation between the measured quantity and the target quantity." [Cochran, et al. 1954, p. 33.]

". . . . . Some people prefer the term 'true value', although others excoriate it as philosophically unsound.

"We could also call the reference level a 'target value'. In a way this is a bad term because it implies that it is something we want to find through the measurement process rather than something we ought to find because, like Mt. Everest, it is there. Unfortunately our desires can influence our notion of what is true, and we can even unconsciously bring the latter into agreement with the former; my use of the term 'target value' is not meant to imply that I think it legitimate to equate what we would like to see with what is there." [Murphy 1961, p. 265.]

d. Concepts of the Precision and Accuracy of a Measurement Process

By the *precision* of a measurement process we mean the degree of mutual agreement characteristic of independent measurements of a single quantity yielded by repeated applications of the process under specified conditions; and by its *accuracy* the degree of agreement of such measurements with the true value of the magnitude of the quantity concerned. In other words, the *accuracy* of a measurement process refers to, and is determined by the degree of conformity *to the truth* that is characteristic of independent measurements of a single quantity produced (or producible) by the repeated applications of the process under specified conditions; whereas its *precision* refers solely to, and is determined solely by the degree of conformity *to each other* characteristic of such measurements, irrespective of whether they tend to be close or far from the truth. Thus, *accuracy* has to do with *closeness to the truth; precision*, only with *closeness together*

This distinction between the meanings of the terms "accuracy" and "precision" as applied to measurement processes and measuring instruments is consistent with the etymological roots of these words. "Etymologically the term 'accurate' has a Latin origin meaning 'to take pains with' and refers to the care bestowed upon a human effort to make such effort what it *ought* to be, and 'accuracy' in common dictionary parlance implies freedom from mistakes or exact conformity to *truth*. 'Precise,' on the other hand, has its origin in a term meaning 'cutoff, brief, concise'; and 'precision' is supposed to imply the property of determinate limitations or being exactly and sharply defined." [Shewhart 1939, p. 124.] Thus one can properly speak of a national, state, or local law as being "precise," but not as being "accurate"—to what *truth* can it conform? On the other hand, if one spoke of a particular translation as being "accurate" this would imply a high degree of fidelity to the original "attained by the exercise of care." Whereas, to speak of it as being "precise," would imply merely that it is unambiguous, without indicating whether it is or is not correct.[7]

In spite of the distinct difference between the etymological meanings of the terms "accuracy" and "precision," they are treated as synonyms in many standard dictionaries; and Merriam-Webster [1942], after drawing the helpful distinctions quoted in the foregoing footnote, promptly topples the structure so carefully built by adding "scrupulous exactness" as an alternative meaning of "precise." Consequently it is not surprising that "There are probably few words as loosely used by scientists as *precision* and *accuracy*.—It is not unusual to find them used interchangeably in scientific writings." [Schrock 1950, p. 10.]

On the other hand, as Shewhart has remarked:

> "Careful writers in the theory of errors, of course, have always insisted that accuracy involves in some way or other the difference between what is observed and what is true, whereas precision involves the concept of reproducibility of what is observed. Thus Laws, writing on electrical measurements, says: [8] 'Every experimenter must form his own estimate of the accuracy, or approach to the absolute truth obtained by the use of his instruments and processes of measurement. He must remember that a high precision, or agreement of the results among themselves, is no indication that the quantity under measurement has been accurately determined.' As another example we may take the following comment from a recent and authoritative treatise on chemical analysis: [9] 'The analyst should form the habit of estimating the probable accuracy of his work. It is a common mistake to confuse accuracy and precision. Accuracy is a measure of the degree of correctness. Precision is a measure of reproducibility in the hands of a given operator.' " [Shewhart 1939, pp. 124–125.]

More recently, Lundell, Hoffman, and their associates at the National Bureau of Standards have reemphasized the importance of the distinction between "precision" and "accuracy":

> "In discussions of chemical analysis, the terms precision and accuracy are often used interchangeably and therefore incorrectly, for precision is a measure of reproducibility, whereas accuracy is a measure of correctness. The analyst is vitally interested in both, for his results must be sufficiently accurate for the purpose in mind, and he cannot achieve accuracy without precision, especially since his reported result is often based on one determination and rarely on more than three determinations. The recipient of the analysis is interested in accuracy alone, and only in accuracy sufficient for his purposes." [Hillebrand et al., 1953, p. 3.]

It is most unfortunate that in everyday parlance we often speak of "accuracy and precision," because *accuracy* requires *precision*, but *precision* does not necessarily imply *accuracy*.

> "It is, in fact, interesting to compare the measurement situation with that of a marksman aiming at a target. We would call him a precise marksman if, in firing a sequence of rounds, he were able to place all his shots in a rather small circle on the target. Any other rifleman unable to group his shots in such a small circle would naturally be regarded as less precise. Most people would accept this characterization whether either rifleman hits the bull's-eye or not.
>
> "Surely all would agree that if our man hits or nearly hits the bull's-eye on all occasions, he should be called an accurate marksman. Unhappily, he may be a very precise marksman, but if his rifle is out of adjustment, perhaps the small circle of shots is centered at a point some distance from the bull's-eye. In that case we might regard him as an inaccurate marksman. Perhaps we should say that he is a potentially accurate marksman firing with a faulty rifle, but speaking categorically, we should have to say that the results were inaccurate." [Murphy 1961, p. 265.]

It follows from what has been said thus far that "if the precisions of two processes are the same but the biases are different, the process of smaller bias may be said to have higher accuracy while if the biases are both negligible, the process of higher precision may be said to have higher accuracy." Unfortunately, "in other cases such a simple comparison may be impossible." [ASTM 1961, p. 1760.]

---

[7] It is sometimes helpful to distinguish between "correct," "accurate," and "exact": "CORRECT, the most colorless term, implies scarcely more than freedom from fault or error, as judged by some (usually) conventional or acknowledged standard; . . . ACCURATE implies, more positively, fidelity to fact or truth attained by the exercise of care; . . . EXACT emphasizes the strictness or rigor of the agreement, which neither exceeds nor falls short of the fact, standard or truth; . . . PRECISE stresses rather sharpness of definition or delimitation . . ." [Merriam-Webster 1942 p. 203].

[8] Frank A. Laws, Electrical Measurements, p. 593 (McGraw-Hill, New York, N.Y., 1917).

[9] G. E. F. Lundell and J. I. Hoffman, Outlines of Methods of Chemical Analysis, p. 220 (John Wiley and Sons, New York, N.Y., 1938).

To fully appreciate the preceding statement—and especially the difficulty of comparing accuracies in some cases—let us consider figures 1 and 2, in which the origins of the scales correspond to the true value of τ of the quantity measured, so that the curves shown may be regarded as depicting the distributions of *errors* of the measurements yielded by a selection of different measurement processes. Consider first the three symmetrical distributions in the top half of figure 1. All three of these distributions are centered on zero, so that these measurement processes have no *bias*. It is evident that the process of highest precision, c, is also the process of highest accuracy; and that the process of least precision, a, is also the process of least accuracy. Since curve b in the upper half of figure 1 and curve d in the lower half have identical size and shape, the corresponding processes have the same *precision;* but process b is without bias, whereas process d has a positive bias of two units, so that process b is clearly the more *accurate*. (In particular we may note that whereas it is practically certain that process b will not yield a measurement deviating from the truth by more than two units, exactly one-half of the measurements yielded by process d will deviate from the truth by this much or more.) Similar remarks clearly apply to processes c and e corresponding to curve c in the upper half and curve e in the lower half of figure 1, but in this instance the superiority of process c relative to process e with respect to *accuracy* is even more marked. (In particular, we may note that whereas it is practically certain that no measurement yielded by process c will deviate from the truth by as much as one unit, it is practically certain that every measurement yielded by process e will deviate from the truth by more than one unit.)

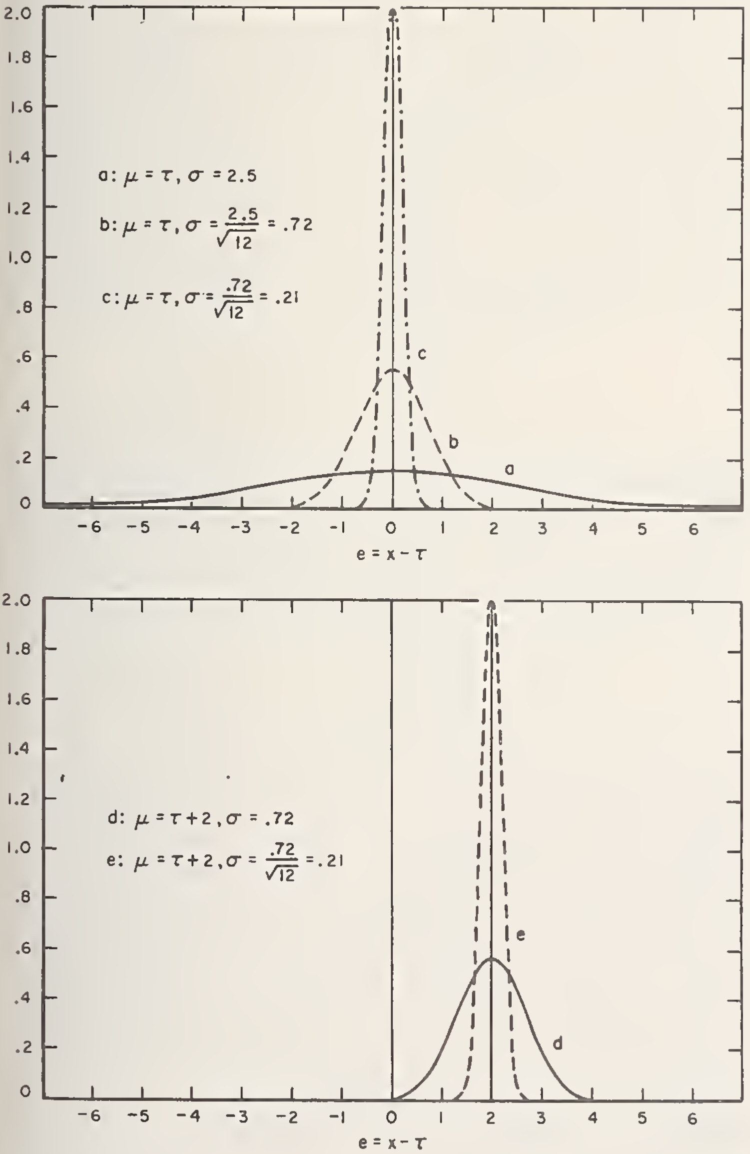


FIGURE 1. *Distributions of errors of some biased and unbiased measurement processes of various precisions.*

Figure 2, which is essentially the same as one given by General Simon [1946, fig. 1], portrays three measurement processes A, B, and C, differing from each other with respect to both precision and bias. Comparison of these three processes with respect to *accuracy* is not quite so simple. First, it is evident that, although process A has greater precision than process B, process B is the more accurate of the two. (In particular, it is practically certain that none of the measurements yielded by process B will deviate from the truth by more than 4 units, whereas 50 percent of the measurements from process A will deviate from the truth by four units or more.) Next, is process B more (or less) *accurate* than process C which is *unbiased*, but has a very low *precision*? Process B has a positive *bias* of two units, but has sufficiently greater *precision* than process C to also have greater *accuracy* than process C. (While approximately 50 percent of the measurements

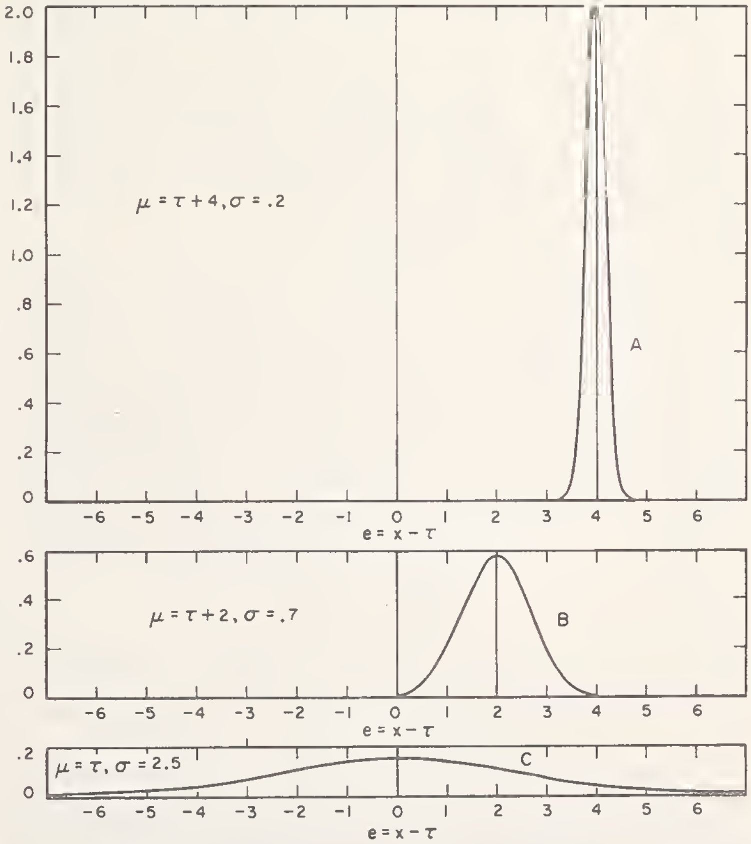


FIGURE 2. *Three measurement processes differing from each other with respect to both precision and accuracy.*

yielded by process C will deviate from the truth by more than two units (in either direction), and exactly 50 percent of the measurements yielded by process B will deviate from the truth by two units or more (in the positive direction only), it cannot be ignored that about 10 percent of the measurements yielded by process C will deviate from the truth by four units or more whereas it is practically certain that no measurement yielded by process B will deviate from the truth by as much as four units.) Similarly, it may be argued that process A, in spite of its bias, has greater *accuracy* than process C "since the range in measurements of C more than covers the corresponding ranges of A or B." [Simon 1946, p. 654.] While this conclusion that of the three measurement processes depicted in figures 2, process C has the least *accuracy*, may not be entirely acceptable to some persons, it is consistent with Gauss' dictum, in a letter to F. W. Bessel, to the effect that maximizing the probability of a zero error is less important than minimizing the "average" injurious effects of errors in general. [C. F. Gauss, 1839, pp. 146–147.]

Before leaving figure 2, we must not fail to join General Simon in remarking that "the average of a large number of measurements from [process] C will be more accurate than a similar average from either A or B" [Simon 1946, p. 654]. This point is actually illustrated in our figure 1: the three curves in the top half of figure 1 portray the distributions of errors of *single* measurements (curve *a*) of *averages of 12* measurements (curve *b*) and *averages of 144* measurements (curve *c*) from process C; and curves *d* and *e* in the lower half show the distributions of errors of *individual* measurements (curve *d*), and of *averages of 12* measurements (curve *e*) from process B, respectively. It is evident that *averages of 12* measurements from process C (curve *b* in upper portion of fig. 1) have not only greater *accuracy* than *individual* measurements from process B (curve *d* in lower portion of the figure), but also greater *accuracy* than *averages of 12* measurements from process B (curve *e* in lower portion).

On the other hand, it is obvious that, if our choice is between individual measurements from process C (curve *a*) and *averages of 12* measurements from process B (curve *e*), the latter will clearly provide greater *accuracy*. In brief, *a procedure with a small bias and a high precision can be more accurate than an unbiased procedure of low precision.* It is important to realize this, for in practical life it is often far better to always be quite close to the true value than to deviate all over the place in individual cases but strictly correct "on the average," like the duck hunter who put one swarm of shot ahead of the duck, and one swarm behind, lost his quarry, but had the dubious satisfaction of knowing that in theory he had hit it "on the average." This we must remember: in practical life we rarely make a very large number of measurements of a given type—we can't wait to be right on the average—our measurements must stand up in individual cases as often as possible.

Despite the foregoing, freedom from bias, that is, freedom from "large" bias, is a desirable characteristic of a measurement process. After all we want our measurements to yield us a determination that we can use as a substitute for the unknown value of a particular magnitude whose value we need for some purpose—we don't want a determination of the value of some other magnitude whose relation to the one we need is indefinitely known.

In view of the difficulty of comparing with respect to *accuracy* measurement processes that differ both in *bias* and *precision*, some writers have elected to take the easy way out by defining "accuracy" to be equivalent to absence of bias, saying that of two measurement processes having different biases, the process of smaller bias is the more "accurate" regardless of the relation of their respective precisions. (See, for example, Beers [1953, p. 4], Ostle [1954, p. 4], and Schenck [1961, p. 4, p. 14].) While the adoption of this concept of "accuracy" certainly makes the discussion of "accuracy" and "precision" simpler for the authors concerned, this practice is contrary to the principle of "conservation of linguistic resources," as R. B. Murphy puts it, adding: "It seems to me that the terms 'bias' and 'systematic error' are adequate to cover the situation with which they are concerned. If, nevertheless, we add the term 'accuracy' to apply again in this restricted sense, we are left wordless—at the moment at least—when it comes to the idea of over-all error. From the point of view of the need for a term it is hard to defend the view that accuracy should concern itself solely with bias. . . . [and] there is overwhelming evidence that we need a term at least for the concept of over-all error." [Murphy 1961, pp. 265–266.]

3.4. Mathematical Specification of the Precision of a Measurement Process

a. Simple Statistical Control

Let us now consider the mathematical definition of the *precision* of a measurement process under a fixed set of circumstances. By definition, the precision of a measurement process has to do with the "closeness together" that is typical of successive measurements of a single quantity generated by applications of the process under these fixed conditions. Otherwise expressed, it has to do with the typical "closeness together" of the two individual measurements constituting an arbitrary pair. If the expression "*typical* 'closeness together'" is to be meaningful, the measurements generated by repeated application of the process to the measurement of a single quantity must be homogeneous in some sense. Therefore, for the moment, let us assume that the measurement process is in a state of *simple statistical control*, so that the successive measuremets in each of the sequences (1), ($i=1, 2, 3, . . .$), generated by the process may *all* be regarded as "observed" values of independent identically distributed random variables.

Just as we may regard each individual measurement x_{ij} in a particular sequence (1) as striving to express the value of the limiting mean μ, so also we may regard each individual difference $x_{ij}-x_{ik}$, $j\neq k$, as striving to express the characteristic spread between an arbitrary pair of measurements, x' and

x'', say. For this purpose the signs of these differences are clearly irrelevant. Therefore, by analogy with our use of a sequence of cumulative arithmetic means, (2), to achieve a mathematical formulation of the concept of a limiting mean associated with measurement of a given quantity by a particular measurement process, let us adopt the sequence of cumulative arithmetic means of the *squares* of the $n(n-1)/2$ distinct differences among the first n measurements of a particular sequence (1), for example, the sequence

$$\overline{(d^2)}_{in} \equiv \frac{2}{n(n-1)} \sum_{j=1}^{n-1} \sum_{k=j+1}^{n} (x_{ij}-x_{ik})^2, \qquad (n=2, 3 \ldots), \quad (3)$$

as the basis of a mathematical formulation of the concept of the precision of a measurement process.

The necessary and sufficient condition for almost sure convergence of the sequence (3) to a finite limit, say Δ^2, is that the Strong Law of Large Numbers be applicable to the sequence.

$$x_{i1}^2, x_{i2}^2, \ldots, x_{ij}^2 \ldots, \quad (4)$$

consisting of the squares of the corresponding terms of the original sequence (1). (Boundedness of the x's in addition to statistical control is, for example, sufficient to ensure that the sequence (4) will also obey the Strong Law of Large Numbers.) If the Strong Law of Large Numbers is applicable to the sequence of squares (4), and if the measurement process is in a state of simple statistical control, then the cumulative arithmetic means of the squares of the measurements, that is, the sequence

$$\overline{(x^2)}_{in} \equiv \sum_{j=1}^{n} x_{ij}^2/n, \qquad (n=1, 2, \ldots), \quad (5)$$

will almost surely tend to a limit, say S, the magnitude of which will depend on the quantity measured, the measurement process involved, but not on the "occasion" (identified by the subscript "i"). By virtue of an algebraic identity that is well known to students of mathematical inequalities, namely,

$$n\sum_{j=1}^{n} a_j^2 - \left(\sum_{j=1}^{n} a_j\right)^2 = \frac{1}{2}\sum_{j=1}^{n}\sum_{k=1}^{n} (a_j-a_k)^2, \qquad (n \geq 2) \quad (6)$$

and of the fact that the right-hand side of (6) is always positive except when the a's are all equal, it is easily seen, on dividing both sides of (6) by n^2, that S will always exceed μ^2, the square of the (almost sure) limit of the sequence (2), so that we may write $S=\mu^2+\sigma^2$, with $\sigma^2>0$. Furthermore, applying the algebraic identity (6) in reverse to the right-hand side of (3) yields the following relationship between the corresponding terms of sequences (3), (5), and (1):

$$\overline{(d^2)}_{in} = 2\left(\frac{n}{n-1}\right)\left\{\overline{(x^2)}_{in} - (\bar{x}_{in})^2\right\} > 0, \qquad (n \geq 2). \quad (7)$$

Hence, if a measurement process is in a state of simple statistical control and the Strong Law of Large Numbers is applicable to a *sequence* of squared measurements (4), then the sequence $\overline{(d^2)}_{in}$, defined by (3), will, in view of (7), tend almost surely to a finite limit $\Delta^2=2\sigma^2$. Thus we see that σ^2, termed the *variance* of the measurement process, is the mean value of one-half of the squared difference between two arbitrary measurements x' and x'', that is,

$$\sigma^2 = \tfrac{1}{2}\overline{(x'-x'')^2}, \quad (8)$$

and provides an indication of the *im*precision of the process. The square root of the variance, σ, is termed the *standard deviation* of the process.

It is natural, therefore, on the basis of a single sequence of n measurements of a single quantity, to take

$$s^2 \equiv \frac{1}{2}\overline{(d^2)} = \frac{1}{n(n-1)} \sum_{j=1}^{n-1}\sum_{k=j+1}^{n} (x_j-x_k)^2 = \frac{\sum_{j=1}^{n}(x_j-\bar{x})^2}{n-1} \quad (9)$$

as the sample estimate of the underlying variance σ^2; and the square root, s, as the sample estimate of σ.[10]

From (9), since $\bar{x} \equiv \bar{x}_n$ tends (almost surely) to μ it is evident that σ^2 is also the mean value of the squared deviations of individual measurements from the limiting mean μ of the process, that is $\sigma^2 = \overline{(x-\mu)^2}$, so that the standard deviation σ may be regarded, in the language of mechanics, as the radius of gyration of the distribution of all possible measurements x about μ, the limiting mean of the process.

Remark: Mathematically the foregoing discussion can be carried out equally well in terms of the absolute (unsigned) values of the differences instead of in terms of their squares. Such an approach is, mathematically speaking, somewhat more general in that it requires for its validity merely that the Strong Law of Large Numbers be applicable to the sequence $|x_{i1}|, |x_{i2}|, \ldots, |x_{ij}|, \ldots$ of *absolute values* of the x_{ij} rather than to the sequence (4) of their squares. From the practical viewpoint, however, this greater generality is entirely illusory, and the mathematics of absolute values of variables is always more cumbersome than the mathematics of their squares. For example, the arithmetic mean of the absolute values of the $n(n-1)/2$ *distinct* differences among n measurements, i.e.,

$$\overline{|d|}_n \equiv \frac{2}{n(n-1)} \sum_{j=1}^{n-1}\sum_{k=j+1}^{n} |x_j-x_k| \quad (10)$$

---

[10] From the algebraic identity (6), it is evident that the practice in some circles of dividing $\sum_{j=1}^{n}(x-\bar{x})^2$ by n, instead of $n-1$, amounts to including each of the *distinct* squared differences $(x_j-x_k)^2$, $j \neq k$, *twice* in the summation, together with n identically zero terms $(x_j-x_k)^2$, $j=k$, each included once, and then dividing by n^2, the total number of terms (real and phantom) involved. Viewed in this light it would seem that division by $n-1$ is more reasonable, in that the inclusion of identically zero terms in the formulation of a measure of *variation* is a bit unreasonable.

is *not* expressible as a multiple of the sum of the absolute deviations of the measurements from their mean, $\sum|x_i-\bar{x}|$, and for large values of n the evaluation of (10) presents computational difficulties. The approach in terms of the absolute values of the differences also has the disadvantage from the practical viewpoint that, as we shall see in a moment, *components of imprecision* are additive in terms of squared quantities such as σ^2, so that in this sense the *variance* σ^2 is a more appropriate measure of the dispersion of the x's about their limiting mean μ than is σ itself.

Ordinarily, the magnitude of σ^2 (and, hence, of σ), unlike that of μ, depends only on the measurement process concerned and the circumstances under which it is applied, and not also on the magnitude of the quantity measured—otherwise we could not speak of a measurement process having *a* variance, or *a* standard deviation.

Since the *precision* of the process obviously decreases as the value of σ (or, of σ^2) increases, and vice versa, it is necessary to take some inverse function of σ as a measure of the precision of process. To conform with traditional usage it is necessary to regard the precision of a measurement process as *inversely proportional* to its *standard deviation* σ which is, therefore, a measure of the *imprecision* of the process. Thus, Gauss, writing in 1809, remarked that his constant $h=1/\sigma\sqrt{2}$ could properly be considered to be a measure of the precision of the observations because if, for example $h'=2h$, that is, if $\sigma'=\frac{1}{2}\sigma$, then "a double error can be committed in the former system with the same facility as a single error in the latter, in which case, according to the common way of speaking, a double degree of precision is attributed to the latter observations." [11]

The fact of the matter is, however, that:

> ". . . different fields have particularly favorite ways of expressing precision. Most of these measures are multiples of the standard deviation; it is not always clear which multiple is meant. . . .
>
> "Some consider it unfortunate that precision should be stated as a multiple of standard deviation, since precision should increase as standard deviation decreases. Indeed, it would be more exact to say that standard deviation is a measure of imprecision. However, sensitivity, as we have previously indicated, suffers from this logical inversion without hurt. Perhaps we can best avoid this by saying that standard deviation is an index of precision. The habit of saying 'The precision is . . . ' is deeply rooted, and there would be understandable impatience with the notion that standard deviation should be numerically inverted before being quoted in a statement of precision." [Murphy 1961, pp. 266–267.]

In consequence the ASTM has, at least tentatively, taken the following position:

> "The numerical value of any commonly used index of precision will be smaller the more closely bunched are the individual measurements of a process. As more causes are added to the system, the greater the numerical value of the index of precision will ordinarily become. If the same index of precision is used on two different processes based on the same method or intended to measure the same physical property, the process that has the smaller value of the index of precision is said to have higher precision. Thus, although the more usual indexes of precision are really direct measures of *imprecision*, this inversion of reference has been firmly established by custom. The value of the selected index of precision of a process is referred to simply as its precision or its stated precision." [ASTM 1961, p. 1759.]

As we have remarked previously, in practical work the end result of measuring some quantity or calibrating an instrument for a standard rarely consists of a single measurement of the quantity of interest. More often it is some kind of average or adjusted value, for example, the arithmetic mean of a number of independent measurements of the quantity of interest. Let us, therefore, consider the statistical properties of a sequence of *arithmetic means* of successive nonoverlapping groups of n measurements each from a sequence (1) of *individual* measurements yielded by a measurement process on a particular occasion. In other words, let us consider the sequence

$$\bar{x}_{i1}, \bar{x}_{i2}, \ldots, \bar{x}_{im}, \ldots \tag{11}$$

of distinct arithmetic means of n measurements each

$$\bar{x}_{im}=\frac{1}{n}\sum_{j=(m-1)n+1}^{mn} x_{ij}, \quad (m=1, 2, \ldots), \tag{12}$$

derived from a sequence (1) of individual measurements of a single quantity produced, or at least conceptually producible, by the measurement process concerned on, say, the ith occasion. If the "underlying measurement process" giving rise to the individual measurements x_{ij} is in a state of simple statistical control, then the "extended measurement process" giving rise to the averages $\bar{x}_{im}$ will also be in a state of simple statistical control. Consequently, the mathematical analysis of section 3.2, but with the averages $\bar{x}_{im}$ in place of the individual measurements x_{ij}, will carry through without other change. Let $\mu_{\bar{x}}$ denote the limiting mean thus associated with the "extended measurement process" giving rise to the averages $\bar{x}_{im}$ as its "individual" measurements. Since the cumulative arithmetic mean of the first m terms of the sequence (11) is the same as the cumulative arithmetic mean of the first mn terms of the sequence (1) of individual measurements, it is clear that the limiting mean $\mu_{\bar{x}}$ associated with the sequence of averages (11) is the same as the limiting mean associated with the original sequence (1) of individual measurements, that is,

$$\mu^{-}=\mu_x=\mu. \tag{13}$$

Similarly, the mathematical analysis at the beginning of the present section, but with the individual measurements x_{ij} in (3) thru (9), replaced by the averages $\bar{x}_{im}$, carries through essentially as before. Let $\sigma_{\bar{x}}^2$ denote the variance thus associated with the "extended measurement process" giving rise to the sequence of averages (11). As in the case of the variance σ^2 of individual measurements,

---

[11] "Ceterum constans h tamquam mensura praecisionis observationum considerari poterit. . . . Quodsi igitur e.g., $h'=2h$, aeque facile in systemate priori error duplex committi poterit, ac simplex in posteriori, in quo casu observationi ibus posterioribus secundum vulgarem loquendi morem praecisio duplex tribuitur." [Gauss 1809, Art. 178; 1871, p. 233; English translation, 1857, pp. 259–260.]

so also may $\sigma^2_{\bar{x}}$ be interpreted as the overall mean value of the squared deviation of "individual" averages $\bar{x}$ from the limiting mean $\mu_{\bar{x}}$ of the "extended process," that is,

$$\sigma^2_{\bar{x}}=\overline{(\bar{x}-\mu_{\bar{x}})^2}=\overline{(\bar{x}-\mu)^2} \tag{14}$$

By virtue of the algebraic identity

$$\begin{aligned}(\bar{x}-\mu)^2&=\left[\frac{1}{n}\sum_{j=1}^{n}x_j-\mu\right]^2=\left[\frac{1}{n}\sum_{j=1}^{n}(x_j-\mu)\right]^2\\ &=\frac{1}{n^2}\left[\sum_{j=1}^{n}(x_j-\mu)^2+2\sum_{j=1}^{n-1}\sum_{k=j+1}^{n}(x_j-\mu)(x_k-\mu)\right]\end{aligned} \tag{15}$$

it is readily seen that

$$\sigma^2_{\bar{x}}=\frac{\sigma^2_x}{n}=\frac{\sigma^2}{n}. \tag{16}$$

(The mean value of a sum is always the sum of the mean values of its individual terms, so that the overall mean value of the first summation inside the brackets in the last line of (15) is simply $n\sigma^2_x$. Furthermore, in the case of independent identically distributed measurements, the overall mean value of the term involving the double summation is 0.)

Since, from (16), $\sigma_{\bar{x}}=\sigma/\sqrt{n}$, it is seen that the precision of the arithmetic mean of n *independent* measurements is proportional to $\sqrt{n}$. Hence the arithmetic mean of *4* independent measurements has *double* the precision of a single measurement; the mean of *9* independent measurements, *thrice* the precision of a single measurement; and *144* independent measurements will be required if their arithmetic mean is to have a *12-fold* increase in precision over a single measurement. (But to ask for a 12-fold increase in precision is to ask for a very considerable improvement indeed, as can be seen from a comparison of curves *a* and *c* in the top half of fig. 1.)

To serve as a reminder of the distinction between the standard deviation of an individual measurement and the standard deviation of a mean $\bar{x}$, it is customary to refer to σ as the "standard *deviation*" of a single measurement x, and to $\sigma_{\bar{x}}$ as the "standard *error*" of the (arithmetic) mean $\bar{x}$.

b. Within-Occasions Control

In the foregoing it has been assumed that the individual measurements comprising the sequences (1) corresponding to the respective "occasions," ($i=1, 2, \ldots$), could *all* be regarded as "observed values" of independent identically distributed random variables, that is, that the measurement process concerned was in a state of simple statistical control. When such is the case then any subset of n measurements is strictly comparable to any other subset of n measurements, and any two such subsets can be combined and regarded validly as a single set of $2n$ measurements. Unfortunately, as Student's comment quoted on page 167 above clearly implies, such complete homogeneity of measurement is rarely if ever met in practice. More often the situation is as described by Sir George Biddell Airy, British Astronomer Royal 1835–1881, in (to my knowledge) the first elementary book on the theory of errors and combination of observations in the English language [Airy 1861, p. 92]:

> "When successive series of observations are made, day after day, of the same measurable quantity, which is either invariable . . . or admits of being reduced by calculation to an invariable quantity . . .; and when every known instrumental correction has been applied . . .; still it will sometimes be found that the result obtained on one day differs from the result obtained on another day by a larger quantity than could have been anticipated. The idea then presents itself, that possibly there has been on some one day, or on every day, some cause, special to the day, which has produced a *Constant Error* in the measures of that day."

Sir George, however, cautions against jumping to conclusions on the basis of only a few observations:

> "The existence of a daily constant error . . . ought not to be lightly assumed. When observations are made on only two or three days, and the number of observations on each day is not extremely great, the mere fact, of accordance on each day and discordance from day to day, is not sufficient to prove a constant error. [And we should interject here that under such circumstances apparent over-all accordance is not sufficient to prove the absence of daily constant errors either.] The existence of an accordance analogous to a 'round of luck' in ordinary changes is sufficiently probable. . . . More extensive experience, however, may give greater confidence to the assumption of constant errors . . . first, it ought, in general to be established that there is possibility of error, constant on one day but varying from day to day. . . ." [Airy 1861, p. 93.]

The most useful statistical tools for this purpose are the control-chart techniques of the industrial quality control engineer. If in such a situation, a series of measurements obtained by measurement of a single quantity a number of times on each of several different days or "occasions" by a particular measurement process is plotted in the form of a *control chart for individuals* [ASTM 1951, pp. 76–78, and pp. 101, 105], the individual measurements so plotted will be seen to consist of "sections" identifiable with the subsequences (1) corresponding to the respective "occasions," ($i=1, 2, 3, \ldots$), with the measurements within sections pair-wise closer together on the average than two measurements one of which comes from one section and the other from another. Such a series of measurements is clearly "out of control." If now parallel $\bar{x}$- and R-charts are constructed from these data, based on a series of samples of equal size from *within* the respective "occasions" or "sections" *only*, i.e., excluding means $\bar{x}$ and ranges R of any samples that "straddle" two occasions, and the points on the resulting $\bar{x}$-chart are clearly "out of control," then we may infer the existence of day-by-day components of error, constant, perhaps, on one day, but varying from day to day.

If points on the R-chart constructed as described are "out of control" also, then the measurement operation concerned is in a completely unstable condition and cannot be described validly as a "measurement process" at all. On the other hand, if the $\bar{x}$-chart is "out of control," but the R-chart is "in control," then we may regard the measurement process as being in a state of *within-occasions control.* ("It is usually not safe to conclude that a state of control exists unless the plotted points for at least 25 successive subgroups fall within the 3-sigma control limits. In addition, if not more than 1 out of 35 successive points, or not more than 2 out of 100, fall outside the 3-sigma control limits, a state of control may ordinarily be assumed to exist." [ASA 1958c, p. 18.]) In such a situation we postulate the existence of (at least, conceptually) different limiting means μ_i for the respective "occasions" $(i=1, 2, . .)$, and a common *within-occasions variance* σ^2_w.

An unbiased estimate of the *within-occasions standard deviation* σ_w can be obtained, if desired, from the average range $\overline{R}$ used in constructing the R-chart, by means of the formula

$$\text{unbiased estimate of } \sigma_w=\overline{R}/d_2 \qquad (17)$$

where d_2 is the factor given in the d_2 column of table B2 of [ASTM 1951, p. 115] corresponding to the sample or subgroup size n used in constructing the R-chart.

Alternatively, if desired, an unbiased estimate of σ^2_w can be obtained directly from the measurements involved by means of the formula

$$\text{unbiased estimate of } \sigma^2_w=s^2_w\frac{\sum_{h=1}^{k}\sum_{j=1}^{n}(x_{hj}-\bar{x}_h)^2}{k(n-1)}, \qquad (18)$$

where x_{hj} denotes the jth measurement and $\bar{x}_h$ the arithmetic mean of the n measurements of the hth subgroup, respectively, and k is the number of subgroups involved in constructing the R-chart.

c. Complex or Multistage Control

When a measurement process is not in a state of simple statistical control that satisfies the criteria of within-occasions control, that is, when the $\bar{x}$-chart (and control chart for individuals) are clearly "out of control," but the 25 or more subgroup ranges plotted on the R-chart exhibit control, then it is usually of importance to ascertain whether the measurement process concerned is possibly in a state of *complex or multistage statistical control.* For this purpose four or more measurements from each of at least 25 different occasions will be needed. Taking one sample of n successive measurements, $(4\leq n\leq 10)$, from the available measurements corresponding to each of, say, $k(\geq 25)$ different "occasions," evaluate the arithmetic means $\bar{x}_i$ of these samples, $(i=1, 2, . . ., k)$, and *treating these averages as INDIVIDUAL measurements* construct a control chart for these "individuals" and parallel $\bar{x}$- and R-charts as described in [ASTM 1951, Example 22, p. 101]. If the points plotted on these three control charts exhibit control, then we "act for the present as if" the measurement process concerned is in a state of *complex or multistage statistical control* and regard the limiting means μ_i for the respective "occasions," $(i=1, 2, . . .)$ as being in a state of simple statistical control with a limiting mean μ and variance σ^2_b, termed the *between-occasions* component of variance.

If in such a situation we were to form cumulative arithmetic means such as (3) of the squares of all distinct differences between arbitrary pairs of measurements from *within* each of the respective "occasions," then such cumulative arithmetic means of squares of differences would almost surely tend to $2\sigma^2_w$ in the limit as the number of pairs included tends to infinity, where σ^2_w is the "within-occasions variance" mentioned above in connection with "within-occasions control." If, on the other hand we were to form similar cumulative arithmetic means of the squares of differences between arbitrary pairs consisting in each instance of one measurement from each of two different sections, then such a cumulative arithmetic mean of squared differences would tend almost certainly to $2(\sigma^2_w+\sigma^2_b)$ as the number of "occasions" sampled tends to infinity, where σ^2_b is the above mentioned "between-occasions variance," i.e., the variance of the limiting means μ_i for the respective "occasions" about their limiting mean μ.

If in utilizing measurements from a measurement process that is in such a state of complex statistical control, one forms an average $\bar{x}_N$ that is the arithmetic mean of a total of $N=kn$ measurements, composed of n measurements from each of k different "occasions," then the variance of $\bar{x}_N$ will be

$$\sigma^2_{\bar{x}_N}\equiv\overline{(x_N-\mu)^2}=\frac{1}{k}\left(\sigma^2_b+\frac{\sigma^2_w}{n}\right) \qquad (19)$$

From (19) it is clear that, if σ^2_b is at all sizable compared to σ^2_w, then, for fixed $N=kn$, $\bar{x}_N$ will have greater precision *as a determination of* μ when based on a large number k of different occasions, with only a small number n of measurements from each occasion. Finally, setting $k=1$, we see that the mean $\bar{x}_i$, of n measurements all taken on the same occasion *considered as a determination of the overall limiting mean* μ has an *overall variance* $\sigma^2_{\bar{x}}=\sigma^2_b+(\sigma^2_w/n)$; but *considered as a determination of* μ_i, the limiting mean for the ith occasion, its variance is only σ^2_w/n. In other words, the "standard error" of a mean such as $\bar{x}_i$ is not unique, but depends on the purpose for which it is to be used.

An unbiased estimate of the overall standard deviation $\sigma_{\bar{x}_i}$ of the arithmetic mean of n measurements taken on a single "occasion" may be obtained by the procedure of formula (17) above, if desired, using the average range $\overline{R}$ employed in constructing the R-chart corresponding to the groups of averages $\bar{x}_{in}$.

Alternatively, an unbiased estimate of the overall variance $\sigma^2_{\bar{x}}$ can be obtained directly from the means $\bar{x}_i$ used in constructing the $\bar{x}$-chart, by using the formula

$$s^2_{\bar{x}}=\frac{\sum_{i=1}^{k}(\bar{x}_i-\bar{\bar{x}})^2}{k-1} \qquad (20)$$

where $\bar{x}_i$ is the arithmetic mean of the n successive observations from the ith "occasion," $(i=1, 2, \ldots, k)$ and $\bar{\bar{x}}$ is the arithmetic mean of these k means.

The foregoing concept of a state of *complex or multistage statistical control* can be extended readily to more complex truly "multistage" situations involving three or more "levels" of random variation.

Finally, it is evident from the foregoing that when a measurement process is in a state of complex or multistage statistical control, then the difference between two individual measurements (or the arithmetic means of n measurements) corresponding to two different "occasions" will include the difference $\mu_i - \mu_{i'}$ between the limiting means corresponding to the two particular occasions involved. In so far as such a comparison is regarded as a unique individual case, the difference $\mu_i - \mu_{i'}$ is a fixed constant and hence a systematic error affecting this comparison. On the other hand, if the difference between these two individual measurements (or these two arithmetic means) is regarded only as a typical instance of the outcomes that might be yielded by the same measurement process on *other* pairs of occasions, then the difference $\mu_i - \mu_{i'}$ may be regarded as a random component having a zero mean and variance $2\sigma_b^2$.

It goes without saying, of course, that if a control-chart analysis of the type described above is undertaken for the purpose of ascertaining whether the process is in a state of complex control, but the points plotted on the $\bar{x}$-chart are clearly "out of control," then the measurement process concerned cannot be regarded as statistically stable from occasion to occasion, and should be used only for *comparative measurement* within-occasions. Even when such a measurement process is used solely for comparative measurement within "occasions," it needs to be shown that comparative measurements or *fixed differences* are in a state of (simple or complex) statistical control, if this measurement process is to be generally valid in any absolute sense. Thus in the case of the thermometer calibration procedure mentioned in section 2.4 above, one needs to examine the results of repeated measurement, occasion after occasion, of the difference between two standard thermometers S_1 and S_2 of proven stability in order to determine whether the process is or is not in a state of simple or complex statistical control.

3.5. Difficulty of Characterizing the Accuracy of a Measurement Process

Unfortunately, there does not exist any single comprehensive measure of the accuracy (or *in*accuracy) of a measurement process (analogous to the standard deviation as a measure of its imprecision) that is really satisfactory. This difficulty stems from the fact that "accuracy," like "true value," seems to be a reasonably definite concept on first thought, but as soon as one attempts to specify exactly what one means by "accuracy" in a particular situation, the concept becomes illusive; and in attempting to resolve the matter one comes face to face, sooner or later, with the question: "Accurate" for what *purpose*?

Gauss, in his second development (1821–1823) of the Method of Least Squares clearly recognized the difficulty of characterizing sharply the "accuracy" of any particular procedure:

> "Quippe quaestio haec per rei naturam aliquid vagi implicat, quod limitibus circumscribi nisi per principium aliquatenus arbitrarium nequit . . . neque demonstrationibus mathematicis decidenda, sed libero tantum arbitrio remittenda."[12] [Gauss 1823, Part I, Art. 6.]

Gauss himself proposed [loc. cit.] that the *mean square error* of a procedure—that is, $\sigma^2+(\mu-\tau)^2$, where σ is its *standard deviation;* and $\mu-\tau$, its *bias*—be used to characterize its accuracy. While *mean square error* is a useful criterion for comparing the relative accuracies of measurement processes differing widely in both precision and bias, it clearly does not "tell the whole story." For example, if one were to adopt the principle that measurement processes having the same mean square error were equally "accurate," then one would be obliged to consider the measurement processes corresponding to the three curves shown in figure 3 as being of equal

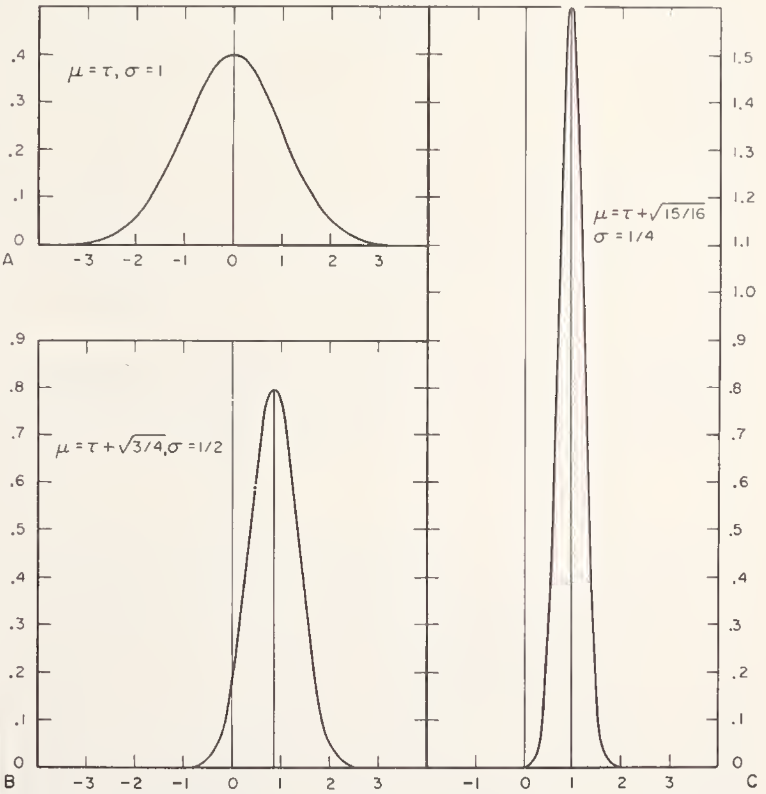


FIGURE 3. *Three distributions differing with respect to both precision and accuracy but with the same mean square error.*

---

[12] I am grateful to my colleague Franz Alt for the following literal translation of these phrases:

"For this question implies, by the very nature of the matter, something vague which cannot be clearly delimited except by somewhat arbitrary principle . . . nor can it be decided by mathematical demonstrations, but must be left to mere arbitrary judgment."

accuracy, whereas for many purposes one would regard process *C* (portrayed to the right) as the "most accurate," in spite of the fact that the chances of scoring a "bull's eye" or "near miss" are greater in the case of process *A* shown in the upper left.

Alternatively, if one were to say that two measurement processes were equally accurate when exactly the same proportion *P* of the measurements of each lay within $\pm\delta$ units from the true value, then for $P=0.5$ one would be obliged to say that the measurement processes corresponding to curves *e* and *d* in the lower half of figure 1 were equally accurate, and that the measurement process corresponding to curve *a* in the upper half of the same figure was slightly more accurate than either *e* or *d*. Or, taking $P=0.95$, one would be obliged to say that the measurement processes corresponding to the three curves shown in figure 4 were equally accurate. From these, and other cases easily constructed, it is readily seen that it is unsatisfactory to regard two measurement processes as being equally accurate if the same specified fraction *P* of the measurements produced by each lie within the same distance from the true value.

Thus one is led by the force of necessity to the inescapable conclusion that ordinarily (at least) two numbers are needed to adequately characterize the accuracy of a measurement process. And this has been recognized by the American Society for Testing and Materials in their recent recommendations [ASTM 1961, pp. 1759–1760]:

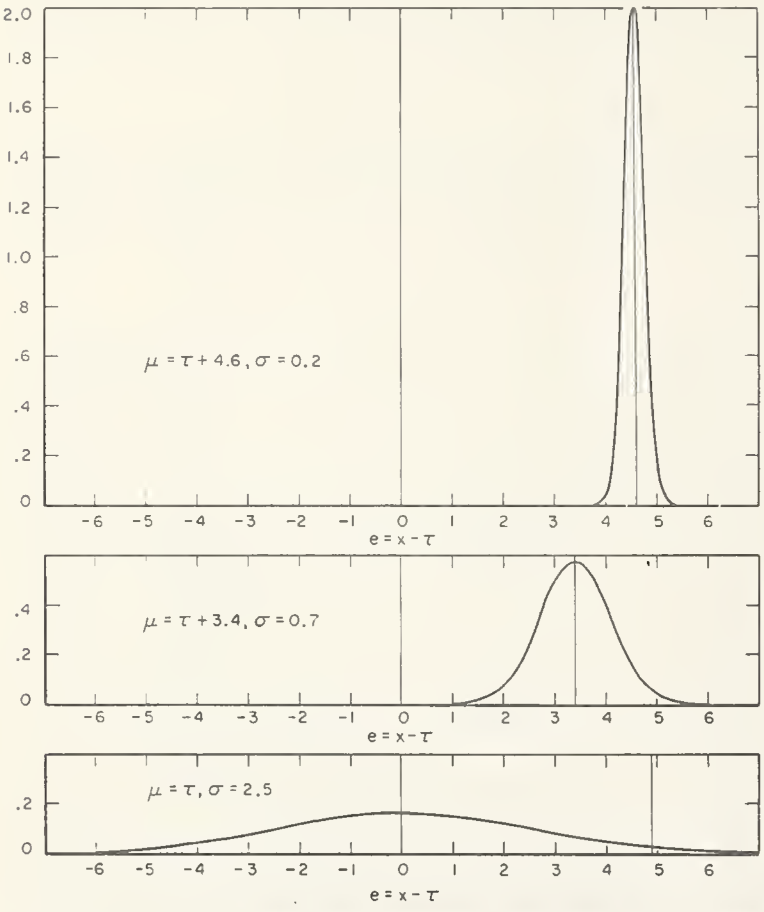


FIGURE 4. *Three measurement processes differing in bias and precision but having 95 percent of their individual measurements within ±4.9 units from the true value τ.*

> "Generally the index of accuracy will consist of two or more different numbers. Since the concept of accuracy embraces not only the concept of precision but also the idea of more or less consistent deviation from the reference level (systematic error or bias), it is preferable to describe accuracy by separate values indicating precision and bias."

The fact of the matter is that *two* numbers ordinarily suffice only because the "end results" of measurement and calibration programs are usually averages or adjusted values based on a number of independent "primary measurements," and such averages and adjusted values tend to be normally distributed to a very good approximation when four or more "primary measurements" are involved. This is illustrated by figure 5, which shows the distributions of individual measurements of two unbiased measurement processes with identical standard deviations but having *uniform* and *normal* "laws of error," respectively, together with the corresponding distributions of arithmetic means of 4 independent measurements from these respective processes—these latter two distributions are depicted by a single curve because the differences between the two distributions concerned are far less than can be resolved on a chart drawn to this scale. Since both of the processes concerned are unbiased, "accuracy" thus becomes only a matter of "precision"—or does it?—both curves for $n=1$ have the same standard deviation, do they reflect equal "accuracy"? Would not the answer depend on the advantages to be gained from small errors balanced against the seriousness of large errors, in relation to the purpose for which a single measurement from one or the other is needed? But "the problem" disappears nicely if averages of 4 measurements are to be used.

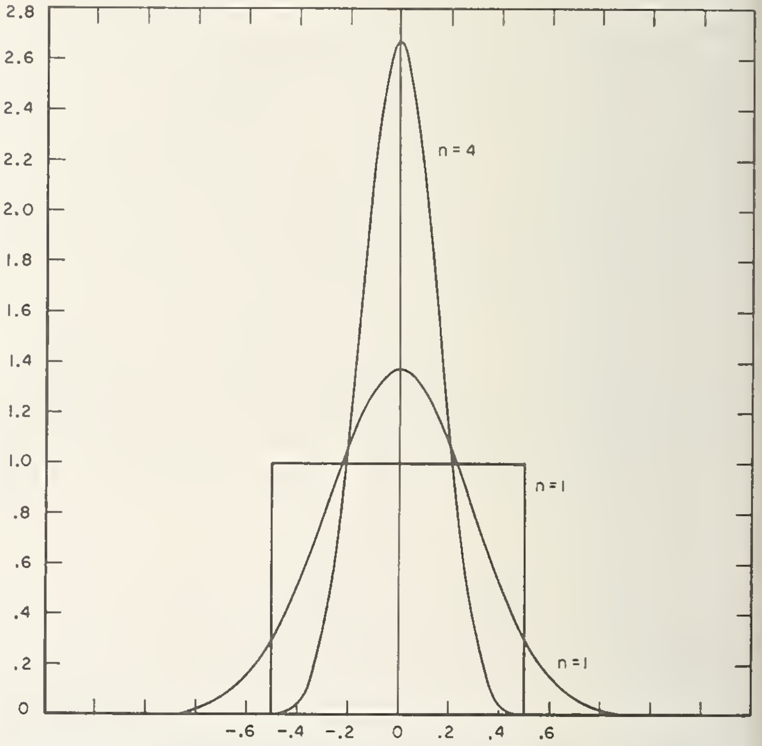


FIGURE 5. *Uniform and normal distributions of individual measurements having the same mean and standard deviation, and the corresponding distribution(s) of arithmetic means of four independent measurements.*

4. Evaluation of the Precision, and of Credible Bounds to the Systematic Error of a Measurement Process

As we have just seen, two numbers are ordinarily needed to characterize the accuracy of a measurement process, the one indicating its *precision*, and the other its *bias*. In practice, however, the bias of a measurement process is unknown and unknowable because the "true values" of quantities measured are almost always unknown and unknowable. The principle exception is when one is measuring a difference that is by hypothesis identically zero. If the bias of a measurement process could be, and were known exactly, then one would of course subtract it off as a "correction" and thus dispose of it entirely. Since ordinarily we cannot expect to know the exact magnitude of the bias of a measurement process, we are forced in practice to settle for credible bounds to its likely magnitude—much as did Steyning and the thief in chapter VI of Kipling's story, *Captains Courageous:* "Steyning tuk him for the reason that the thief tuk the hot stove—bekaze for there was nothing else that season". Consequently, neither the *bias* nor the *accuracy* of any measurement process, or method of measurement, can ever be known in a logical sense. The *precision* of a measurement process, however, can be measured and known. (Compare Deming [1950, p. 17].)

4.1. Evaluation of the Precision of a Measurement Process

In the foregoing we have stressed that a measurement operation to qualify as a *measurement process* must have attained a state of statistical control; and that until a measurement operation has been "debugged" to the extent that it has attained a state of statistical control, it cannot be regarded in any logical sense as measuring anything at all. It is also clear, from our discussion of the control-chart techniques for determining whether in any given instance one is entitled to "act for the present as if" a state of statistical control has been attained, that a fairly large amount of experience with a particular measurement process is needed before one can resolve the question in the affirmative. Once a measurement process has attained a state of statistical control, and so long as it remains in this state, then an estimate of the *standard deviation* of the process can be obtained from the data employed in establishing control, as we have indicated above.

Since the precision of a measurement process refers to, and is determined by the characteristic "closeness together" of successive independent measurements of a single magnitude generated by repeated application of the process under specified conditions, it is clearly necessary in determining whether a measurement operation is or is not in a state of statistical control, and in evaluating its precision to be reasonably definite on what variations of procedure, apparatus, environmental conditions, observers, operators, etc., are allowable in "repeated applications" of what will be considered to be the same measurement process applied to the measurement of the same quantity under the same conditions. If whatever measure of the precision and bounds to the bias of the measurement process we may adopt are to provide a realistic indication of the accuracy of this process in practice, then the "allowable variations" must be of sufficient scope to bracket the range of circumstances commonly met in practice. Scientists and engineers commonly append "probable errors" or "standard errors" to the results of their experiments and tests. These measures of imprecision are supposed to indicate the extent of the reproducibility of these experiments or tests under "essentially the same conditions," but there are great doubts whether the "probable errors" and "standard errors" generally presented actually have this meaning. The fault in most cases is not with the statistical formulas and procedures used to compute such probable errors or standard errors from the measurements in hand, but rather with the limited scope of the "conditions" sampled in taking the measurements.

a. Concept of a "Repetition" of a Measurement

As a very minimum, a "repetition" of a measurement by the same *measurement process* should "leave the door open" to, and in no way inhibit changes of the sort that would occur if, on termination of a given series of measurements, the data sheets were stolen and the experimenter were to repeat the series as closely as possible with the same apparatus and auxiliary equipment following the same instructions. In contrast, a "repetition" by the same *method of measurement* should permit and in no way inhibit the natural occurrence of such changes as will occur if the experimenter were to mail to a friend complete details of the apparatus, auxiliary equipment, and experimental procedure employed—i.e., the written text specification that defines the "method of measurement" concerned—and the friend, using apparatus and auxiliary equipment of the same kind, and following the procedural instuctions received to the best of his ability, were then, after a little practice, to attempt a repetition of the measurement of the same quantity. Such are the extremes, but there is a "gray region" between in which there is not to be found a sharp line of demarcation between the "areas" corresponding to "repetition" by the *same measurement process*, and and to "repetition" by the *same method of measurement*.

Let us consider "repetitions" by the same *measurement process* more fully. Such repetitions will undoubtedly be carried out in the same place, i.e., in the same laboratory, because if it is to be the same measurement process, the very same apparatus must be used. But a "repetition" cannot be carried out at the same *time*. How great a lapse of time should be allowed, nay *required*, between "repetitions"? This is a crucial question. Student gives an answer in a passage from which we quoted above [Student 1917, p. 415]:

> "Perhaps I may be permitted to restate my opinion as to the best way of judging the accuracy of physical or chemical determinations.
>
> "After considerable experience I have not encountered any determination which is not influenced by the date on which it is made; from this it follows that a number of determinations of the same thing made on the same day are likely to lie more closely together than if the repetitions had been made on different days.
>
> "It also follows that if the probable error is calculated from a number of observations made close together in point of time, much of the secular error will be left out and for general use the probable error will be too small.
>
> "Where then the materials are sufficiently stable it is well to run a number of determinations on the same material through any series of routine determinations which have to be made, spreading them over the whole period."

Another important question is: Are "repetitions" by the same measurement process, to be limited to repetitions by the same observers and operators, using the same auxiliary equipment (bottles of reagents, etc.); or enlarged to include repetitions with nominally equivalent auxiliary equipment, by various but equivalently trained observers and operators? I believe that everyone will agree that substitution, and certainly replacement, of bottles of reagents, of batteries as sources of electrical energy, etc., by "nominally equivalent materials" must be allowed. And any calibration laboratory having a large amount of "business" will certainly, in the long run at any rate, have to face up to allowing changes, even replacement of observers and operators—and, ultimately, even of apparatus.

A very crucial question, not always faced squarely, is: in complete "repetitions" by the same measurement process, are such "repetitions" to be limited to those intervals of time over which the apparatus is used "as is" and "undisturbed," or extended to include the additional variations that almost always manifest themselves when the apparatus is disassembled, cleaned, reassembled, and readjusted? Unless such disassembly, cleaning, reassembly, and readjustment of apparatus is permitted among the allowable variations affecting a "repetition" by the same measurement process, then there is very little hope of achieving satisfactory agreement between two or more measurement processes in the same laboratory that differ only in their identification with different pieces of apparatus of the same kind. In practice it is found that statistical control can be attained and maintained under such a broad concept of "repetition" only through the use of reference standards of proven stability. Furthermore, by thus more squarely facing the issue of the scope of variations allowable with respect to "repetitions" by the same measurement process, we shall go a long way toward narrowing the gap between a "repetition" by the same measurement process and by the same method of measurement.

As we have said before, if whatever measures of the precision and bias of a measurement process we may adopt are to provide a realistic indication of the accuracy of this process in practice, then the "allowable variations" must be of sufficient scope to bracket the range of circumstances commonly met in practice. Furthermore, any experimental program that aims to determine the precision and systematic error, and thence the accuracy of a measurement process, must be based on an appropriate random sampling of this "range of circumstances," if the usual tools of statistical analysis are to be strictly applicable. Or as Student put it, "the experiments must be capable of being considered to be a *random* sample of the population to which the conclusions are to be applied. Neglect of this rule has led to the estimate of the value of statistics which is expressed in the crescendo 'lies, damned lies, statistics'." [Student 1926, p. 711.]

When adequate random sampling of the appropriate "range of circumstances" is not feasible, or even possible, then it is necessary to compute, by extrapolation from available data, a more or less subjective estimate of the "precision" of the end results of a measurement operation, to serve as a substitute for a direct experimental measure of their "reproducibility." Youden [1962d] calls this "approach the 'paper-way' of obtaining an estimate of the [precision]." Its validity, if any, "is based on subject-matter knowledge and skill, general information, and intuition—but not on statistical methodology" [Cochran et al. 1953, p. 693].

b. Some Examples of Realistic "Repetitions"

As Student remarked [1917, p. 415], "The best way of judging the accuracy of physical or chemical determination . . . [when] the materials are sufficiently stable . . . is . . . to run a number of determinations on the same material thru any series of routine determinations which have to be made, spreading them over the whole period." To this end, as well as to provide an overall check on procedure, on the stability of reference standards, and to guard against mistakes, it is common practice in many calibration procedures, to utilize two or more reference standards as part of the regular calibration procedure.

The calibration procedure for *liquid-in-glass thermometers*, referred to in section 2.4 above, is a case in point. A measurement of the difference between the two standards S_1 and S_2 is obtained as by-product of the calibration of the four test thermometers T_1, T_2, T_3, and T_4 in terms of the (corrected) readings of the two standards. It is such remeasurements of the difference between a pair of standard thermometers from "occasion" to "occasion" that constitutes realistic "repetitions" of the calibration procedure. The data yielded by these "repetitions" are of exactly the type needed (a) to ascertain whether or not the process is in a state of statistical control; and if so, (b) to determine its overall standard deviation.

Similarly, in the calibration of *laboratory standards of mass* at the National Bureau of Standards, "known standard weights are calibrated side-by-side with [the] unknown weights" [Almer et al., 1962, p. 33]. Indeed, weights whose values are otherwise determined "are not said to have been 'calibrated'. That term is reserved for measurements based on at least two mass standards." [loc. cit., p. 43.] In the specimen work sheets exhibited by Almer et al., the auxiliary standards involved are those from the Bureau's "NH series" of reference standards known by the designations NH50, NH20, and $NH10_1$ respectively. It is the measurements obtained in routine calibrations of the differences between the values of these standards and their accepted values that not only provide valuable checks on day-to-day procedure, but also serve as the basis for determination of the overall standard deviation of this calibration process.

A third example is provided by the method followed at the National Bureau of Standards for testing *alternating-current watthour meters*, which has been described in some detail by Spinks and Zapf [1954]. Four reference watthour meters are involved. One of these, termed "the Standard Watthour Meter," is located in the device portrayed in figure 1 of the paper by Spinks and Zapf. The other three are located in a temperature-controlled cabinet. A "test" of a watthour meter sent to the Bureau involves not only a comparison of this watthour meter with the Standard Watthour Meter, but also comparisons of each of the Comparison Standard Watthour Meters with the Standard Watthour Meter. It is from the data yielded by these intercomparisons of the Standard Watthour Meter and the Comparision Standard Watthour Meters that the standard deviation of this test procedure is evaluated. Spinks and Zapf's section on "Precision and Accuracy Attainable" is notable for its exceptional lucidity as well as for its completeness with respect to relevant details.

Some additional examples of realistic "repetitions" are discussed by Youden [1962c].

4.2. Treatment of Inaccuracy Due to Systematic Errors of Assignable Origins but of Unknown Magnitudes

As we remarked in section 3.3b above, the systematic error of a measurement process will ordinarily have both constant and variable components. For convenience of exposition, it is customary to regard the individual components of the overall systematic error of a measurement or calibration process as elemental or constituent "systematic errors" and to refer to them simply as "systematic errors," for short. Included among such "systematic errors" affecting a particular measurement or calibration process are: ". . . all those errors which cannot be regarded as fortuitous, as partaking of the nature of chance. They are characteristic of the system involved in the work; they may arise from errors in theory or in standards, from imperfections in the apparatus or in the observer, from false assumptions, etc. To them, the statistical theory of error does not apply." [Dorsey 1944, p. 6; Dorsey and Eisenhart 1953, p. 104.]

The overall systematic error of a measurement process ordinarily consists of elemental "systematic errors" due to both assignable and unassignable causes. Those of unknown (not thought of, not yet identified, or as yet undiscovered) origin are always to be feared; allowances can be made only for those of recognized origin.

Since the "known" systematic errors affecting a measurement process ascribable to specific origins are ordinarily determinate in origin only, their individual values ordinarily being unknown both with respect to sign and magnitude, it is not possible to evaluate their algebraic sum and thereby arrive at a value for the overall systematic error of the measurement process concerned. In consequence, it is necessary to arrive at bounds for each of the individual components of systematic error that may be expected to yield nonnegligible contributions, and then from these bounds arrive at credible bounds to their combined effect on the measurement process concerned. Both of these steps are fraught with difficulties.

Determination of reasonable bounds to the systematic error likely to be contributed by a particular origin or assignable cause necessarily involves an element of judgment, and the limits cannot be set in exactitude. By assigning ridiculously wide limits, one could be practically certain that the actual error due to a particular cause would never lie outside of these limits. But such limits are not likely to be very helpful. The narrower the range between the assigned limits, the greater the uneasiness one feels that the assigned limits will not include whatever systematic error is contributed by the cause in question. But a decision has to be made; and on the basis of theory, other related measurements, a careful study of the situation in hand, especially its sensitivity to small changes in the factor concerned, and so forth, "the experimenter presently will feel justified in saying that he feels, or believes, or is of the opinion," that the systematic error due to the particular source in question does not exceed such and such limits, "meaning thereby, since he makes no claim to omniscience, that he has found no reason for believing" that it exceeds these limits. In other words, 'nothing has come to light in the course of the work to indicate" that the systematic error concerned lies outside the stated range. [Dorsey 1944, pp. 9–10; Dorsey and Eisenhart, 1953, pp. 105–107.]

This being done to each of the recognized potential sources of systematic error, the problem remains how to determine credible bounds to their combined effect. Before considering this problem in detail, it will be helpful to digress for a moment, to consider an instructive example relating to the combined effect of constant errors in an everyday situation.

a. An Instructive Example

Consider the hypothetical situation of an individual who is comparing his checkbook balance with his bank statement. To this end he needs to know the total value of his checks outstanding. Loathing addition, or perhaps, simply to save time, he adds up only the dollars, neglecting the cents, and thus arrives at a total of, say, $312, for 20 checks outstanding. Adding a correction of 50 cents per check, or $10 in all, he takes $322 as his estimate. Within what limits should he consider the error of this estimate to lie?

The round-off error cannot exceed ±50 cents per check, so that barring mistakes in addition, he can be absolutely certain that the total error of his estimate does not exceed ±$10. But these are extremely pessimistic limits: they correspond to every check being in error by the maximum possible amount and all in the same direction. (Actually the maximum possible positive error is 49 cents per check or +$9.80 in all.)

To be conservative, but not so pessimistic, one might "allow" a maximum error of ±50 cents per check, but consider it reasonable to regard their signs as being equally likely to be plus or minus. In this way one would be led to conclude "with probability 0.95" that the total error lies between ±$7.00; or "with probability 0.99," between ±$8.00, as shown in the column headed "binomial" in table 1, for $n=20$. The "saving" by this procedure is clearly not great.

Table 1. *Limits of error of a sum of n items indicated by various methods of evaluation*

n	Absolute ±	Binomial		Uniform		Triangular		Normal, $2\sigma=0.5$		Normal, $3\sigma=0.5$	
		0.95 ±	0.99 ±	0.95 ±	0.99 ±	0.95 ±	0.99 ±	0.95 ±	0.99 ±	0.95 ±	0.99 ±
1	0.50	0.50	0.50	0.48	0.50	0.39	0.45	0.49	0.64	0.33	0.43
2	1.00	1.00	1.00	0.78	0.90	0.56	0.71	0.69	0.91	0.46	0.61
3	1.50	1.50	1.50	0.97	1.19	0.69	0.88	0.85	1.12	0.57	0.74
4	2.00	2.00	2.00	1.12	1.41	0.80	1.03	0.98	1.29	0.65	0.86
5	2.50	2.50	2.50	1.25	1.60	0.89	1.15	1.10	1.44	0.73	0.96
6	3.00	2.50	3.00	1.38	1.76	0.98	1.29	1.20	1.58	0.80	1.05
7	3.50	3.00	3.50	1.49	1.91	1.06	1.39	1.30	1.70	0.86	1.14
8	4.00	3.50	3.50	1.59	2.05	1.13	1.49	1.39	1.82	0.92	1.21
9	4.50	3.50	4.00	1.69	2.18	1.20	1.58	1.47	1.93	0.98	1.29
10	5.00	4.00	4.50	1.78	2.31	1.26	1.66	1.55	2.04	1.03	1.36
15	7.50	5.50	6.00	2.19	2.88	1.55	2.04	1.90	2.49	1.27	1.69
20	10.00	7.00	8.00	2.53	3.33	1.79	2.35	2.19	2.88	1.46	1.92
25	12.50	8.50	9.50	2.83	3.72	2.00	2.63	2.45	3.22	1.63	2.15
30	15.00	10.00	11.00	3.07	4.03	2.19	2.88	2.68	3.53	1.79	2.35
40	20.00	13.00	14.00	3.58	4.70	2.53	3.33	3.10	4.07	2.07	2.72
50	25.00	16.00	17.00	4.00	5.26	2.83	3.72	3.46	4.55	2.31	3.04
60	30.00	19.00	20.00	4.38	5.76	3.10	4.07	3.80	4.99	2.53	3.33

Alternatively, one might consider it to be more "realistic" to regard the individual errors as independently and uniformly distributed between −50 cents and +50 cents, concluding "with probability 0.95" that the total error does not exceed ±$2.53; or "with probability 0.99," is not greater then ±$3.33—as shown in the columns under the heading "uniform" in table 1. It is clear that a considerable reduction in the estimate of the total error is achieved by this approach.

Strictly speaking, the foregoing analyses via the theory of probability are both inapplicable to the problem at hand: each round-off error is a fixed number between ±50 cents, and their sum is a fixed number between ±$10. If it were true that round-off errors in such cases were uniformly distributed between ±50 cents, then, if one made a habit of evaluating limits of error according to this procedure, one could expect the limits of error so calculated to include the true total error in 95 percent, or 99 percent of the instances in which this procedure was used in the long run. Round-off errors in such cases are almost certainly not uniformly distributed between ±50 cents. (Many items are priced these days at $2.98 etc., and this will distort the distribution of the cents-portion of one's bills but added sales taxes no doubt have a "smoothing" effect.)

Nevertheless, I believe that you will agree that if, in the hypothetical case under discussion, the checkbook balance, with an allowance of $322 for checks outstanding, failed to agree with the bank statement to within $2.53 (or $3.33), our "friend" would do well to check into the matter more thoroughly. And, alternatively, if his checkbook balance so adjusted, and the bank statement, agreed to within $2.53 (or $3.33), it would be reasonably "safe" for him to "act for the present as if" his balance and the bank statement were in agreement. (See Eisenhart [1947a, p. 218] for discussion of a similar example relating to computation with logarithms.)

b. Combination of Allowances for Systematic Errors

The foregoing example suggests that a similar procedure be used for arriving at credible limits to the likely overall effect of systematic errors due to a number of different origins. A number of additional difficulties confront us, however, in this case. To begin with, in view of the inexactness with which bounds can ordinarily be placed on each of the individual components of systematic error, it is not possible to say with absolute certainty that their combined effect lies between the sum of the positive bounds and the sum of the negative bounds.

Second, even if it were possible to scale the situation so that the bounds for each of the components of systematic error was the same, say, $\pm\Delta$, there would still remain the problem of translation into an appropriate probability calculus. Most persons would, I believe, regard the "binomial" approach (corresponding to equal probability of maximum error in either direction), as too pessimistic; and the approach via a uniform distribution of error, as a bit conservative, on the grounds that one intuitively feels that the individual errors are somewhat more likely to lie near the centers than near the ends of their respective ranges. Therefore, one might attempt to simulate this "feeling" by assuming the "law of error" to be an isosceles triangle centered at zero and ends at $\pm\Delta$; or, more daringly, by assuming the "law of error" to be approximately normal with Δ corresponding to 2 "σ" or even 3 "σ."

Unfortunately whatever "probability limits" may be placed upon the combined effects of several independent systematic errors by these procedures are quite sensitive to the assumption made at this stage, as is evident from table 1. Therefore, anyone who uses one of these methods for the "combination of errors" should indicate explicitly which of these (or an alternative method) he has used. When (a) the number of systematic errors to be combined is large, (b) the respective ranges are approximately equal in size, and (c) one feels "fairly sure" that the individual errors do not fall outside of their respective ranges, then my personal feeling is that the "uniform" method is probably a wee bit conservative but "safe"; the triangular method is a bit "too daring"; the normal method with "σ"$=\Delta/3$ ordinarily "much too daring"; but the normal method with "σ"$=\Delta/2$, probably "not too daring." When (b) and (c) hold but n is small, then it will probably be safe to use the "uniform" method with "Δ" taken equal to the average of the individual ranges. Other cases, e.g., when n is large but, say, one or two of the ranges is (are) much larger than the others and tend(s) to dominate the situation, requires special consideration which is beyond the scope of the present paper.

4.3. Expression of the Inaccuracy of a Measurement Process

By whatever means credible bounds to the likely overall systematic error of the measurement process are obtained they should not be combined (by simple addition, by "quadrature," or otherwise) with an experimentally determined measure of its standard deviation to obtain an overall index of its accuracy (or, more correctly, of its *in*accuracy). Rather (a) the standard deviation of the process and (b) credible bounds to its systematic error should be stated separately, because, as we showed in figure 3, a measurement process having standard deviation $\sigma=0.25$ and a bias $\Delta=\sqrt{15/16}=0.97$ is for most purposes "more accurate" than a measurement process having zero bias and standard deviation $\sigma=1$, so that a process with $\sigma=0.25$ and a bias *less than* ±0.97 will *a fortiori* be "more accurate."

Finally, if the uncertainties in the assigned value of a national standard or of some fundamental constant of nature (e.g., in the *volt as maintained at the National Bureau of Standards*, or in the speed of light c, or in the acceleration of gravity g on the Potsdam basis) is an important potential source of systematic error affecting the measurement process, no allowance for possible systematic error from this source should be included ordinarily in evaluating overall bounds to the systematic error of the measurement process. Since the error concerned, what ever it is, affects all results obtained by the method of measurement involved, to include an allowance for this error would be to make everybody's results appear unduly inaccurate relative to each other. Instead, in such instances one should state (a) that results obtained by the measurement process concerned are in terms of the volt (or the watthour, or the kilogram, etc.) "as maintained at the National Bureau of Standards" [McNish and Cameron 1960, p. 102], or "correspond to the speed of light $c=2.997925\times10^{10}$ cm/sec. *exactly*," say; and (b) that the indicated bounds to the systematic error of the process are exclusive of whatever errors may be present from this (or these) source(s). Given such information, experts can make such additional allowances, as may be needed, in fundamental scientific work; and comparative measurements within science and industry within the United States will not appear to be less accurate than they very likely are for the purposes for which they are to be used.

It is a pleasure to acknowledge the technical assistance of Janace A. Speckman in several phases of the preparation of this paper.

5. Bibliography

Airy, George Biddell (1861), On the Algebraical and Numerical Theory of Errors of Observations and the Combination of Observations (Macmillan and Co., Cambridge and London).

Almer, H. E., L. B. Macurdy, H. S. Peiser, and E. A. Weck (1962), Weight calibration schemes for two knife-edge direct-reading balances, J. Research NBS **66C** (Eng. and Instr.) No. 1, pp. 33–44.

American Standards Association (1958a), Guide for quality control, American Standard Z 1.1–1958. (American Standards Association, 70 East Forty-fifth St., New York 17, N.Y.)

American Standards Association (1958b), Control chart method of analyzing data, American Standard Z 1.2–1958. (American Standards Association, 70 East Forty-fifth St., New York 17, N.Y.)

American Standards Association (1958c), Control chart method of controlling quality during production. American Standard Z 1.3–1958. (American Standards Association, 70 East Forty-fifth St., New York 17, N.Y.)

American Society for Testing Materials (1951), ASTM Manual on Quality Control of Materials Special Technical Publication **15–C** (American Society for Testing Materials 1916 Race St., Philadelphia 3).

American Society for Testing and Materials (1961), Use of the terms precision and accuracy as applied to measurement of a property of a material, ASTM Designation: E 177–61T. Reprinted from ASTM Standards, Pt 11, pp. 1758–1766.

Baird, D.C. (1962), Experimentation: An Introduction to Measurement Theory and Experiment Design, (Prentice-Hall, Inc., Englewood Cliffs, N.J.).

Beers, Yardley (1953), Introduction to the Theory of Error, (Addison-Wesley Publishing Co., Cambridge 42, Mass.).

Bicking, Charles A. (1952), The reliability of measured values—an illustrative example. Photogrammetric Engineering **XVIII**, pp. 554–558.

Cameron, J. M. (1951), The use of components of variance in preparing schedules for the sampling of baled wool, Biometrics **7**, pp. 83–96.

Chauvenet, William (1868), A Manual of Spherical and Practical Astronomy Vol. II, 4th edition, (J. B. Lippincott and Co., Philadelphia).

Cochran, William G., Frederick Mosteller, and John W. Tukey (1953), Statistical problems of the Kinsey report, J. Am. Stat. Assoc. **48**, pp. 673–716.

Cochran, William G., Frederick Mosteller, and John W. Tukey (1954), Principles of sampling, J. Am. Stat. Assoc. **49**, pp. 13–35.

Crow, Edwin L. (1960), An analysis of the accumulated error in a hierarchy of calibrations. IRE Trans. Instr. **I–9**, pp. 105–114.

Deming, W. Edwards and Raymond T. Birge (1937), On the Statistical Theory of Errors, reprinted from Reviews of Modern Physics **6,** pp. 119–161 (1934) with additional notes dated 1937 (The Graduate School, U.S. Department of Agriculture, Washington 25, D.C.).

Deming, W. Edwards (1943), Statistical Adjustment of Data (John Wiley & Sons, Inc., New York, N.Y.)

Deming, W. Edwards (1950), Some Theory of Sampling (John Wiley & Sons, New York, N.Y.).

Dorsey, N. Ernest (1944), The velocity of light, Transactions American Philosophical Society **XXXIV,** pp. 1–110.

Dorsey, N. Ernest and Churchill Eisenhart (1953), On absolute measurement, The Scientific Monthly **LXXVII,** pp. 103–109.

Eisenhart, Churchill (1947a), Effects of rounding or grouping data, Chapter 4 of Techniques of Statistical Analysis, edited by C. Eisenhart, M. W. Hastay, W. A. Wallis (McGraw-Hill Book Co., New York, N.Y.).

Eisenhart, Churchill (1947b), Planning and interpreting experiments for comparing two standard deviations, Chapter 8 of Techniques of Statistical Analysis, edited by C. Eisenhart, M. W. Hastay, W. A. Wallis (McGraw-Hill Book Co., New York, N.Y.).

Eisenhart, Churchill (1949), Probability center lines for standard deviation and range charts, Industrial Quality Control **vi,** pp. 24–26.

Eisenhart, Churchill (1952), The reliability of measured values—fundamental concepts, Photogrammetric Engineering **XVIII,** pp. 542–554 and 558–565.

Eisenhart, Churchill (1962), On the realistic measurement of precision and accuracy, ISA Proceedings of the Eight National Aero-Space Instrumentation Symposium held in Washington, May 1962, pp. 75–83.

Feller, William (1957), An Introduction to Probability Theory and its Applications, Vol. 1, 2d edition (John Wiley & Sons, New York, N.Y.).

Galilei, Galileo (1638), Discorsi e Dimostrazioni Matematiche Intorno a Due Nuove Scienze, Leiden.

Galilei, Galileo (1898), Discorsi e Dimostrazioni Matematiche Intorno a Due Nuove Scienze, Le Opere di Galileo Galilei (Edizione Nazionale) **VIII,** pp 39–448, Firenze.

Galilei, Galileo (1914), Dialogues Concerning Two New Sciences, translated by Henry Crew and Alfonso de Salvio, with an Introduction by Antonio Favaro (The Macmillan Co., New York, N.Y.).

Gauss, C. F. (1809), Theoria Matus Corporum Coelestium in Sectionibus Conicis Solem Ambientium, Frid. Perthes et I. H. Besser, Hamburg; reprinted in Carl Friedrich Gauss Werke, Band **VII,** Gotha, 1871.

Gauss, C. F. (1823), Theoria Combinationis Observationum Erroribus Minimis Obnoxiae Commentationes societatis regiae scientiarum Gottingensis recentiores, **V,** pp. 1–104, Gottingae; reprinted in Carl Friedrich Gauss Werke, Band **IV,** Göttingen, 1873.

Gauss, C. F. (1839), letter to F. W. Bessel dated February 28, 1839, reproduced in "Kritische bemerkungen zur methode der kleinsten quadrate," pp. 142–148 in Carl Friedrich Gauss Werke, Band **VIII,** (B. G. Teubner, Leipzig, 1900).

Gauss, C. F. (1857), Theory of the Motion of the Heavenly Bodies Moving About the Sun in Conic Sections; English translation by Charles Henry Davis (Little, Brown and Co., Boston).

Gnedenko, B. V. (1962), The Theory of Probability (English translation by B. D. Sechler), (Chelsea Publishing Co., New York, N.Y.).

Hermach, F. L. (1961), An analysis of errors in the calibration of electric instruments, Communication and Electronics (AIEE) **54,** pp. 90–95.

Hillebrand, W. F., G. E. F. Lundell, H. A. Bright, J. I. Hoffman, Applied Inorganic Analysis, 2d ed. (1953), (John Wiley & Sons, Inc., New York, N.Y.).

Holman, Silas Whitcomb (1892), Discussion of the Precision of Measurements, (John Wiley and Sons, New York, N.Y.).

Keyser, Cassius J. (1922), Mathematical Philosophy, (E. P. Dutton and Co., New York, N.Y.).

Kline, S. J. and F. A. McClintock (1953), Describing uncertainties in single-sample experiments, Mech. Eng. **75,** pp. 3–8.

Laplace, Pierre Simon (1886), Theorie Analytique Des Probabilites; 3d edition, Vol. 7 of Oeuvres Completes de Laplace publiées sous les auspices de l'Académie des Sciences, (Gauthier-Villars, Imprimeur-Libraire de l'École Polytechnique, du Bureau des Longtudes, Successeur de Mallet-Bachelier, Quai des Grands-Augustins, 55, Paris)

McNish, A. G. and J. M. Cameron (1960), Propagation of error in a chain of standards, IRE Trans. Instr. **I–9,** pp. 101–104.

Millikan, R. A. (1903), Mechanics, Molecular Physics, and Heat, (Ginn and Co., New York, pp. 195–196).

Murphy, R. B. (1961), On the meaning of precision and accuracy, Materials Research and Standards **4,** pp. 264–267.

NPL (1957), Calibration of temperature measuring instruments, National Physical Laboratory Notes on Applied Science, No. 12, pp. 29–30, (Her Majesty's Stationery Office, London).

Ostle, Bernard (1954), Statistics in Research, (The Iowa State College Press, Ames, Iowa).

Parzen, Emanuel (1960), Modern Probability Theory and its Applications, (John Wiley & Sons, New York, N.Y.).

Proschan, Frank (1953), Confidence and tolerance intervals for the normal distribution, J. Am. Stat. Assoc. **48,** pp. 550–564.

Rossini, F. D. and W. Edwards Deming (1939), The assignment of uncertainties to the data of chemistry and physics, with specific recommendations for thermochemistry, J. Wash. Acad. Sci. **29,** pp. 416–441.

Schenck, Hilbert, Jr. (1961), Theories of Engineering Experimentation, (McGraw-Hill Book Co., Inc., New York, N.Y.).

Schrock, Edward M. (1950), Quality Control and Statistical Methods, (Reinhold Publishing Corp., New York 18, N.Y.).

Shewhart, W. A. (1931), Economic Control of Quality of Manufactured Product, (D. Van Nostrand Company, Inc., New York, N.Y.).

Shewhart, Walter A. (1939), Statistical Method from the Viewpoint of Quality Control, (The Graduate School, U.S. Department of Agriculture, Washington, D.C.).

Shewhart, Walter A. (1941), Contribution of statistics to the science of engineering, University of Pennsylvania Bicentennial Conference, Volume on Fluid Mechanics and Statistical Methods in Engineering, pp. 97–124, (University of Pennsylvania Press, Philadelphia).

Simon, Leslie E. (1941), An Engineer's Manual of Statistical Methods, (John Wiley & Sons, Inc., New York, N.Y.).

Simon, Leslie E. (1942), Application of statistical methods to ordnance engineering. J. Am. Stat. Assoc. **37,** pp. 313–324.

Simon, Leslie E. (1946), On the relation of instrumentation to quality control, Instruments **19,** pp. 654–656 (Nov. 1946); reprinted in Photogrammetric Engineering **XVIII,** pp. 566–573 (June 1952).

Spinks, A. W. and T. L. Zapf (1954), Precise comparison method of testing alternating-current watthour meters. J. Research NBS **53,** pp. 95–105.

Student (1908), The probable error of a mean, Biometrika **VI,** No. 1, pp. 1–25.

Student (1926), Mathematics and agronomy, Journal of the American Society of Agronomy **XVIII,** 703–719.

Student (1927), Errors of routine analysis. Biometrika, **XIX,** pp. 151–164.

Swindells, James F. (1959), Calibration of liquid-in-glass thermometers, NBS Circ. 600, pp. 11–12, (U.S. Government Printing Office, Washington 25, D.C.).

Tukey, J. W. (1960), Conclusions vs. decisions, Technometrics **2,** No. 4, pp. 423–433.

Waidner, C. W. and H. C. Dickinson (1907), On the standard scale of temperature in the interval 0° to 100° C, Bul. Bur. Stds. **3,** pp. 663–728.

Webster's Dictionary of Synonyms (1942, 1st ed.), (G. and C. Merriam Co., Springfield, Mass.).

Youden, W. J. (1950), Comparative tests in a single laboratory, ASTM Bulletin No. 166, pp. 48–51.

Youden, W. J. and J. M. Cameron (1950), Use of statistics to determine precision of test methods, Symposium on Application of Statistics, Special Technical Publication No. 103, pp. 27–34, (American Society for Testing Materials, Philadelphia).

Youden, W. J. (1951a), Statistical Methods for Chemists, (John Wiley & Sons, New York, N.Y.).

Youden, W. J. (1951b), Locating sources of variability in a process, Ind. Eng. Chem. **43,** pp. 2059–2062.

Youden, W. J. (1953), Sets of three measurements: The Scientific Monthly **LXXVII**, pp. 143–147.

Youden, W. J. (1954–1959), Statistical Design, Industrial and Engineering Chemistry, Feb. 1954 to Dec. 1959, Bimonthly articles collected in a single booklet available from Reprint Department, ACS Applied Publications, 1155 Sixteenth St., Washington 6, D.C. ,

Youden, W. J., W. S. Connor, and N. C. Severo (1959), Measurements made by matching with known standards, Technometrics **1,** pp. 101–109.

Youden, W. J. (1960), The sample, the procedure, and the laboratory, Anal. Chem. **32,** pp. 23A–37A.

Youden, W. J. (1961a) How to evaluate accuracy, Mat. Res. & Std. **1,** pp. 268–271.

Youden, W. J. (1961b), What is the best value? J. Wash. Acad. of Sci. **51,** pp. 95–97.

Youden, W. J. (1961c), Statistical problems arising in the establishment of physical standards, Proceedings Fourth Berkeley Symposium on Mathematical Statistics and Probability **III,** pp. 321–335 (University of California Press, Berkeley and Los Angeles.).

Youden, W. J. (1961d), Systematic errors in physical constants, Phys. Today **14,** pp. 32–42 (1961); also in Technometrics **4,** pp. 111–123 (1962).

Youden, W. J. (1962a), Experimentation and Measurement, National Science Teachers Association Vistas of Science Series No. 2, (Scholastic Book Services, New York 36, N.Y.).

Youden, W. J. (1962b), Measurement Agreement Comparisons, presented at the Standards Laboratory Conference, National Bureau of Standards, Boulder, Colo., August 8–10, 1962.

Youden, W. J. (1962c), Uncertainties in calibration, IRE Trans. Instr. **I–11,** pp. 133–138 (1962).

Youden, W. J. (1962d), Realistic estimates of errors in measurements, ISA Journal **9,** No. 10, pp. 57–58.

NBS Miscellaneous Publication 248
Proceedings of the 1962 Standards Laboratory Conference

Session 2. Error Analysis of Measurement Systems

Paper 2.2. A Mathematical Approach to the Determination of Calibration System Measurement Error

C. D. Ferris *

(Abstract Only)

Errors in a calibration system may be of random or fixed origin. Various mathematical techniques can be applied to separate these errors. Once the nature of the errors has been determined, various procedures exist for analyzing the experimental data such that one ultimately arrives at a realistic error value for a calibration system. Various considerations include examination of the system statistics when a large number of determinations of system behavior has been performed (treatment of data), possible indirect interaction of system errors-dependent, independent, and correlated errors. On this basis, one may now develop a mathematical expression which indicates total error in terms of individual system errors.

Frequently the experimenter is tempted to derive a confidence level for a given measurement based upon his error analysis. This technique is basically fallacious and an illustrative example is presented. Generally the formulation of a confidence level leads to unsure and unrealistic results.

An example is presented which indicates the propagation of a calibration error in a chain of standards. Another example includes a treatment of the combination of errors for a simple experiment.

---

*Center for Measurement Science, The George Washington University, Washington, D. C.

Session 2. Error Analysis of Measurement Systems

Paper 2.3. Error Analysis in Metrology

D. B. Schneider *

The growing need of more reliable measurements necessitates the ability to analyze measurement system results for their total uncertainties. A measurement result to which no confidence level may be assigned has no definite interpretation. A preliminary survey of the literature has disclosed no clear-cut mathematical solution to this problem. Therefore, an attempt has been made to devise a logical and practical solution.

This paper will present an analytical method which has been developed and is presently employed at the Boeing Primary Standards Laboratory. The method recognizes and combines quantitative values for systematic and random component errors which significantly influence the measurement results. Included in this method is a technique determining the confidence interval of uncertainty.

An example of the application of this technique will be illustrated for the calibration of an optical pyrometer.

1. Introduction

This paper is written to aid a person who is primarily concerned with the analysis of the error in the calibration of instruments. The discussion should not be taken as a definitive, or even, in many respects, generally accepted, treatment, but it is hoped that it will be valuable to other standards personnel as a concrete attempt to face the problem of analyzing the errors always present in their work.

2. Essential Concepts

When a person reads an instrument he does not usually obtain the value of the quantity measured. If the observations of an instrument are to be meaningful, an analysis of the error associated with the measurement is necessary. Before the analysis of error may be accomplished several essential concepts must be established.

The most important concept to be developed is that of an error. Any deviation of the value of the quantity measured from the value taken to represent the measurement of this quantity is the error in that measurement.

There are two types of error: systematic and random. A systematic error will cause the value of a measurement to be displaced a fixed amount from the value of the quantity measured; therefore, a systematic error is a constant error. A random error has random causes and will cause several independent measurements of a quantity to be distributed randomly about their average value.

Figure 1 is presented to aid in the understanding of following concepts and definitions.

(1) T is postulated to be a true value which, if known, would completely determine the quantity measured.

(2) X_i represents the value of the *ith* independent measurement of several measurements made upon the value of T. Henceforth we shall call such a measurement an observation.

---

*Formerly with the Boeing Company, Seattle, Wash. Present address: Lockhead Missiles & Space Co., Sunnyvale, Calif.

(3) n is the number of observations made.

(4) $\overline{X}$ is the average value of all observations and is computed from the formula.

$$\overline{X} = \frac{1}{n} \sum_{i=1}^{n} X_i \tag{1}$$

(5) s is the standard deviation of the observations and is defined by the equation

$$s = \left[\frac{1}{(n-1)} \sum_{i=1}^{n} (X_i - \overline{X})^2 \right]^{\frac{1}{2}} \tag{2}$$

The standard deviation is an important statistic as it may be taken as a measure of the amount of random error affecting the observations. The larger the spread of values about $\overline{X}$ the greater is the standard deviation and the greater use was the effect of the random error.

(6) μ (mu) is the limiting mean of the observations. Because the average of a random error taken over many independent repetitions of the same measurement will approach zero (in a probability sense) as the number of measurements increase, $\overline{X}$ will approach some definite value. The value that $\overline{X}$ approaches as the observations increase without limit is μ the limiting mean.

(7) E_s is the systematic error in the measurement, and as seen from figure 1, is given by the equation

$$E_s = \mu - T \tag{3}$$

(8) E_T is the total error taking $\overline{X}$ as the best estimate of T ; mathematically

$$E_T = \overline{X} - T = (\mu - T) + (\overline{X} - \mu) \tag{4}$$

(9) L, the limit of an error E, is a number such that $-L \leq E \leq L$. Thus, if the limit of the systematic error E_s is L_s then $-L_s \leq E_s \leq L_s$.

A reasonable upper limit to the random error $\overline{X} - \mu$ in eq. (4) is given by $st/\sqrt{n}$, where t is found in tables of Student's distribution^2 and depends on the ''confidence coefficient'' selected and the number of degrees of freedom of the standard deviation (n-1 in this case). (See reference 1 for background.)

If limits of error L_s and $st/\sqrt{n}$ have been obtained for the two errors on the right of eq. (4), there still remains the question of how to combine them to obtain a limit to the total error of $\overline{X}$ as an estimate of T. Also if systematic errors are known to arise from several different causes, there is the same question of how to combine their limits of error.

Two methods commonly used to add limits of error are the quadrature and linear methods of addition. Quadrature addition is correctly used when one can assume that all of the errors added are uncorrelated. If it is known that the errors are correlated then the linear method is used. Linear addition consists of adding all limits of error together whereas quadrature addition squares each limit and takes the square root of the sum of the squares to represent the total limit of error.

The quadrature method gives smaller total limits than does the linear method of addition. The linear method gives an upper bound in any case.

In this case it is not unreasonable to use the quadrature method of addition. Thus we can express our evaluation of T in terms of the estimate $\overline{X}$ and the limits of total error as

$$T = \overline{X} \pm \sqrt{L_s^2 + s^2 t^2 / n} \tag{5}$$

Before this equation is of use, it will be necessary to estimate L_s, the limit of the error E_s .

Before the limit of a systematic error can be discovered it must be expected that the error exists. One may never know if all the systematic errors affecting a measurement are known, but one may lower the chances of an unknown systematic error when the function of each component within a measurement system is known.

2.1. Generalized Measurement System

Figure 2 depicts the general measurement system. By way of description the source contains a quantity upon which the measurement is performed, the sensor is designed to respond to a quantity proportional to the quantity measured, the converter converts the stimulus from the sensor into a stimulus acceptable to the indicator, the indicator changes the stimulus from the converter into a form of intelligence perceptible to the observer, and the observer interprets the result presented by the indicator.

2.2 Measurement Errors

It is possible to inspect a measurement system for possible unknown systematic errors. All errors with which the author is familiar may be grouped into several classifications. The following are errors which could be present in any measurement performed.

(1) Alignment error, E_{AL} , is caused when the relative position of the source to the sensor is such that the true value is not accurately sensed.

Suppose a radiometer is used to measure the normal steradiance from a radiating surface which is known to obey Lambert's cosine law. It cannot be known that the angle between the radiometers line of sight and the normal to the surface is zero. If this angle is not zero, then an alignment error will exist. Thus the possibility of a very small angle causes the possibility of an alignment error.

The normal steradiance, the quantity measured by the instrument, is not changed by the existence of this angle. It just is not sensed.

L_{AL}, the limit of the alinement error, can be computed by estimating the limitations to the angle, computing the difference between the normal steradiance and the steradiance at the maximum possible value of this angle, and equating this difference to L_{AL} .

(2) Calibration error, E_C, is the deviation of the true value of a standard from the nominal value of the standard used.

Every calibrated instrument has a calibration error. This is the error that could exist within the limit of error upon the calibration.

Suppose that a gage block is sent to the National Bureau of Standards and its length is certified to be $L \pm L_C$. Then L_C is the limit of error on E_C , the calibration error for that gage block. The calibration error is systematic in all applications of the gage block.

(3) Interference error, E_{IN} , is caused by interference between an external object and the measurement system component.

If two instruments, A and B, are being used in the same laboratory and it is found that the observations made with A are consistently different when B is on from the observations made with A when B is off, it may then be assumed that B is interfering with A or with the quantity that A measures. The error may be completely eliminated by removing B.

When conditions do not allow for the removal of the interfering object then a limit of error must be determined for this interference error.

(4) Observer error, E_D, is the systematic disagreement between observers concerning an indication.

In systems employing meters and/or brightness or color matching mechanisms it may be found upon test that several observers following the same procedure under similar conditions will obtain results which disagree relative to one another. This disagreement is consistent over many repetitions of the measurement.

The limit L_D to the observer error E_D can be computed by performing a measurement under specific conditions using many observers randomly chosen in several repetitions of the measurement. The mean value of all measurements is computed and three times the standard deviation of the sample members from the mean is used as a measure of L_D .

Thus any observer making a measurement under the same conditions that L_D was determined could have results in error by as much as $\pm L_D$.

(5) Sensor error, E_{SN} , is caused when the sensor acts upon the source to change the value measured.

Because it is impossible to measure a quantity without changing its value, any measurement will have sensor error. The resistance added to a circuit by an ammeter, the cooling or heating of a surface because of heat flow through thermocouple leads, and the heating of the environment by the measurement current in a resistance thermometer are all examples of phenomena which will cause sensor error.

In each of the examples mentioned above the actual value of the quantity measured is changed by the sensor of the instrument performing the measurement.

(6) Source error, E_{SO} , is caused when the conditions necessary to realize the standard source cannot be realized.

Suppose that a lamp which is to serve as a brightness temperature standard is sent to the NBS for calibration and is returned with a table of brightness temperature versus current relationship. In order to realize a specific brightness temperature, one must be able to realize the current value given on the certificate and the ambient temperature at time of calibration and must operate the lamp in a vertical position. But one cannot measure the current exactly. The limit of the error in the current measurement when multiplied by the slope of the temperature versus current curve for the lamp at a specific value of current will give L^{E}_{SO} the limit on the source error E^{E}_{SO} for the lamp for a corresponding brightness temperature due to current measurement error. Similarly, the limits on the source error due to temperature and position should be evaluated. Total source error is a combination of all three.

(7) Transmission error, E_T, is caused when the medium between two components acts upon the intelligence being transmitted.

Any time intelligence in any form is passed from one point to another through a medium, the medium will act to change the intelligence. For example, suppose a radiometer is used to measure the energy from a tungsten cavity and that this energy must pass through a quartz window before falling upon the aperture of the radiometer. The energy measured by the radiometer must be divided by the transmission factor of the quartz. The uncertainty in our knowledge of the value of the transmission will result in an uncertainty in the measured value wherein the transmission error could exist.

The errors classified above may be random or systematic. If they are random, their influence will cause the standard deviation of several observations to increase and, as can be seen from equation (5), the $st/\sqrt{n}$ statistic will increase to take the error into account. If the error is systematic, only a close inspection of the measurement can disclose its existence. The classifications listed above should be used primarily to locate systematic errors.

The errors in any one of the several classifications will usually be independent of the errors in any other classification and thus quadrature addition may be used to add their respective limits of error. It may be found that several errors of one type may be present in a measurement and that these errors are dependent. If this is the case, linear addition should be used to add their respective limits of error. If the errors of one classification depend on the errors in another classification the total limit of error for both classifications should be computed by linear addition.

It is not necessarily true that all of the errors above listed will significantly affect any one measurement; however, it will be assumed for the remainder of this paper that at least one error from each classification effects the measurement and that the total error in one classification is independent from the total error in any other classification.

2.3. Readability

The true indicated value of an observation is the value which would be read from an indicator should the observer make no error. The value of the observation is the approximation that the observer makes of the true indicated value.

The readability (RD) of an indicator is the least interval about the value of the observation wherein it is known that the true indicated value must lie.

Therefore any observation has a readability error of up to ± RD/2 in its representation of the true indicated value.

2.4. Resolution

The resolution (RES) of an instrument is the least change in a given direction in the input which will give a perceptible change in the output in a corresponding direction. A change in the true indicated value of an observation of less than ± RES is not seen. Thus, there is a resolution error of up to ± RES in any true indicated value.

2.5. Limit of Indicator Error

The limit of total error in any one observation caused by the indicator is the limit of error in the true indicated value, RES, plus the limit of error in the observation, RD/2. Thus if the limit of indicator error is represented by L_I, then

$$L_I = (\mathrm{RES} + \mathrm{RD}/2).^{1} \qquad (6)$$

2.6. Total Limit of Systematic Error

The limits of the alinement, calibration, interference, observer, sensor, source, and transmission errors may be approximated. When the limit on each error is known, we may then compute a limit of total error. Therefore, if L_s is the limit of total error E_s, and L_{AL} is the limit of E_{AL}, etc., then L_s is given by the equation

$$L_s = \sqrt{L_{AL}^2 + L_C^2 + L_{IN}^2 + L_D^2 + L_{SN}^2 + L_{SO}^2 + L_T^2 + L_I^2} \qquad (7)$$

3. Calibration Methods

Most calibration methods can be grouped according to identifying characteristics into one of three calibration methods: the direct comparison, the indirect comparison, and the substitution methods. We shall consider each of the above mentioned methods.

3.1. Direct Comparison Method

A source and a measurement instrument comprise the direct comparison method of calibration. Figure 3 shows the possible arrangements of the components.

A non-standard and a standard system appear in each arrangement. The test system is always calibrated by directly comparing the known value of the standard to the measured value of the standard. When these two values do not agree, the test system must be used with a calibration correction term.

3.2. Indirect Comparison Method

A non-standard source and a standard and non-standard instrument comprise the indirect comparison method. The arrangement is indicated in figure 4.

Both the standard and test system measure the value of the source. When the two values do not agree the test system is assumed to be in error and an appropriate correction term is determined.

3.3. Substitution Method

A non-standard source, a non-standard measurement instrument, a standard system, and a test system comprise the substitution method. See figure 5.

With the test system connected, the source is adjusted to some stable value. The measurement systems result is recorded. The standard system is substituted and adjusted to give the same measurement result as was obtained when the test system was connected.

3.4. Calibration

A calibration is performed upon an instrument to find how much the instrument under test disagrees with some accepted standard. If we let $\bar{X}s$ stand for the measurement of a true value as performed by the standard and $\bar{X}$ stand for the measurement of the same true value as performed by a test system; then E_C^T, the deviation of the test measurement from the standard measurement, is given by

$$\bar{X}_s - \bar{X} = E_C^T \pm L_C^T \qquad (8)$$

where L_C^T is the limit of error in E_C^T. If we know E_C^T, we may apply a correction term, equal to E_C^T, so that

$$\bar{X} + C_C^T = \bar{X}_s \pm L_C^T \qquad (9)$$

Thus when C_C^T is applied to the test system, the standard and the test systems will agree to within $\pm L_C^T$.

The following discussion will use the superscripts *S* and *T* for "standard" and "test" respectively. Thus, L_s^T denotes the limit of the systematic error of the test instrument and L_s^S denotes the limit of the systematic error for the standard. Also X_{ij}^S and X_{ij}^T will be used to represent the *ith* observation of the *jth* value made by the standard and test system respectively.

---

[1] A referee questions whether RES should be added to RD/2 to obtain L_I, maintaining it is already included in RD/2 by definition. I do not agree that the error due to resolution is accounted for by the limit to the readability error. The true value could change by as much as the RES before a perceptible change is noticed in the true indicated value. The value of the observation is in error by as much as RD/2 independent of the possible change which could occur within ± RES.—Author

4. Calibration Procedures and Equations

For the sake of brevity the procedure and equations will be given only for the direct comparison method. The procedure for the direct comparison method (See fig. 3a) is given below:

(1) Align the test system to measure the source.

(2) Adjust the standard source to X^S_i.

(3) Record the value of the source X^T_{i1}, , as measured by the test system.

(4) Disconnect or misaline the test system.

(5) Repeat steps one through four n times increasing i one each time.

(6) Change the value of the source from X^S_{i1} to X^S_{i2}.

(7) Perform steps one through five.

(8) Continue the above procedure as many times as are necessary to calibrate the test system.

Analysis of data will be indicated for the first value of the source only. The difference $X^S_{i1} - X^T_{i1}$ is computed for the n times that the measurement of X^S_{i1} is made. The calibration correction term C^T_C is given by

$$C^T_C = \frac{1}{n}\sum_{i=1}^{n} C^T_{C_i}, \qquad (10)$$

where

$$C^T_{C_i} = X^S_{i1} - X^T_{i1}. \qquad (11)$$

Thus,

$$\bar{X}^T_1 + C^T_C = \bar{X}^S_1 \pm L^T_C, \qquad (12)$$

where L^T_C, the limit to the calibration error, is given by

$$L^T_C = \left[(L_s)^2 + \frac{t^2}{n(n-1)} \sum_{i=1}^{n} \left(C^T_{C_i} - C^T_C \right)^2 \right] \qquad (13)$$

and

$$\bar{X}^T_1 = \frac{1}{n}\sum_{i=1}^{n} X^T_{i_1} \quad \text{and} \quad \bar{X}^S_1 = \frac{1}{n}\sum_{i=1}^{n} X^S_{i1} \qquad (14)$$

5. Example

An example of the application of the preceding principles of error analysis follows.

5.1. Statement of Problem

Determine the uncertainty in the calibration of an optical pyrometerat a specific brightness temperature; e.g., 1000°C, using as a standard a tungsten ribbon filament lamp calibrated by the National Bureau of Standards.

5.2. General Discussion

Calibration of the test pyrometer will be by the direct comparison method per figure 3a and the accompanying procedure and equations.

The relationship of the various systematic errors to the measurement system are shown in figure 6 where the errors are represented by symbols within the circles. An S or T superscript denotes standard or test and refers to the lamp or optical pyrometer respectively. The subscripts are as given in the earlier definitions of measurement errors.

The data obtained for the test point in figure 3a, 1000 °C, are listed in table 1.

The following discussion will show how the various systematic errors and the statistical uncertainty are quantitatively evaluated for this example.

5.3. Standard Description

The standard system of figure 6 is shown as a schematic diagram in figure 7. The NBS certificate states that the tungsten ribbon filament lamp will have a brightness temperature of 1000°C when a current of 5.39 amperes is passing through it while it is operated in an upright position in an ambient temperature of 25 °C.

A 0 to 18V, 0 to 50 ampere d-c power supply provides the lamp current I. The value of R is 0.01 ohms to within a calibration error limit of ± 5x10-7 ohms. The voltage across R is measured with a potentiometer calibrated accurate to within $\pm(0.0001V + 2\text{x}10^{-6}$ v) where V is the voltage measured by the potentiometer. The potentiometer has a least count of 10^{-6} v and the limit to the readability is estimated to be equal to the least count. Thus RD/2, the limit to the readability error is $5\text{x}10^{-7}$ v.

TABLE 1: Calibration data

i	BRIGHT to dark	DARK to bright	X^T_{ij}
1	1842	1837	1839.5
2	1842	1836	1839.0
3	1842	1836	1839.0
4	1842	1835	1838.5
5	1841	1837	1839.0
Column	1	2	3

5.4. Evaluation of Systematic Errors

Source Error, E_{SO}^{S}. From the definition of source error and the conditions specified on the NBS certificate, we see that the total source error will be composed of three components: (1) an electrical error E_{SO}^{E} caused by the error in the electrical components, (2) a positional error E_{SO}^{P} caused by deviation from the specified upright position of the lamp, and (3) a thermal error, E_{SO}^{t} caused by a deviation in the ambient temperature of the laboratory in which the calibration is performed from the NBS ambient temperature of 25 °C.

Electrical Source Error, E_{SO}^{E}. In order to realize the standard of 1000 °C one must be able to determine that the current is 5.39 amp. There are several reasons why this can not be accomplished exactly. The calibration uncertainty in R and the calibration and indicator uncertainty in the potentiometer causes a possible systematic error in the determination of I.

The possible error in a voltage determination ΔV is given by

$$\Delta V = \pm \left(L_{IN}^{P} + L_{C}^{P}\right) \tag{15}$$

where L_{IN}^{P} and L_{C}^{P} are the limits of the indicator and calibration error of the potentiometer, respectively.

The terms in equation (15) are given by

$$L_{\mathrm{IN}}^{P} = (\mathrm{RD}/_{2} + \mathrm{RES}) \tag{16}$$

and

$$L_{C}^{P} = 0.0001\mathrm{V} + 2 \times 10^{-6} \text{ v.} \tag{17}$$

For a readability equal to 10^{-6} v, eq.(16) becomes

$$L_{\mathrm{IN}}^{P} = 5 \times 10^{-7} \text{ v.} \tag{18}$$

Substitution of eqs. (17) and (18) into eq. (15) gives

$$\Delta V = \pm (0.0001\mathrm{V} + 2.5 \times 10^{-6}). \tag{19}$$

For a nominal value of $I =$ 5.39 amp. and $R = 10^{-2}$ ohms, we have from ohm's law that

$$V = IR = 5.39 \times 10^{-2} \text{ v.} \tag{20}$$

Substitution of eq. (20) into eq. (19) gives

$$\Delta V = \pm 7.89 \times 10^{-6} \text{ v.} \tag{21}$$

Thus we see that the true value of the voltage across R is equal to the measured voltage plus or minus 7.89×10^{-6} v.

The total error limit in $R, \Delta R$, is the calibration limit of error of 5×10^{-7} ohms,

$$\Delta R = 5 \times 10^{-7}. \tag{22}$$

From Ohm's law we have

$$I = V/R. \tag{23}$$

Therefore, the differential of I is

$$dI = \frac{1}{R} dV - \frac{V}{R^2} dR. \tag{24}$$

Setting $dI = \Delta I$, $dV = \Delta V$ and $dR = \Delta R$ we have

$$\Delta I = \frac{1}{R} \Delta V - \frac{V}{R^2} \Delta R. \tag{25}$$

Substituting eqs. (21) and (22) into eq. (25) yields

$$\Delta I = \pm \frac{(7.89 \times 10^{-6} \text{v})}{\mathrm{R}} \pm \frac{V}{R^2} (5 \times 10^{-7} \text{ ohms}) \tag{26}$$

For $R = 0.01$ ohms and $V = 5.39 \times 10^{-2}$ v, eq. (26) gives

$$\Delta I = \pm 10.58 \times 10^{-4} \text{ amp.} \tag{27}$$

(using linear addition of error limits, though quadrature addition may be justified).

Equation (27) is the uncertainty in the source error in terms of current. A curve was fitted by an electronic computer to the data on the NBS certificate for the lamp. The data in the table below was taken from this curve.

Brightness	Current
°C	amp.
990	5.34
1000	5.39
1010	5.43

From this data, m the slope of the curve is computed as

$$m = (225 \pm 25)\ °\mathrm{C/amp.} \tag{28}$$

Thus, to be conservative, we let

$$m = 250\ °\mathrm{C/amp.} \tag{29}$$

If L_{SO}^{E} represents the maximum source error due to the electrical components then,

$$L_{\mathrm{SO}}^{E} = m \, \Delta I. \tag{30}$$

Substitution of eqs. (27) and (29) into (30) gives

$$L_{\mathrm{SO}}^{E} = 0.265\ °\mathrm{C}. \tag{31}$$

Equation (31) tells us that the source could be in systematic error by as much as ±0.265 C because we could not determine the current exactly.

Positional Source Error, $L_{\mathrm{SO}}^{\mathrm{P}}$. If the lamp is not operated in a vertical position, the heat correction because of gases within the envelope of the lamp will change and cause a corresponding change in the brightness temperature. It has been found from experience that when a plumb bob is used to establish a vertical reference, that no alinement will be off in excess of 1°. The temperature difference corresponding to 1°. is ± 0.50. Thus if E_{SO}^{P} is the source error due to deviation in the vertical position of the lamp,

$$L^{P}_{SO} = \pm 0.50 \text{ °C.} \qquad (32)$$

Thermal Source Error, E^{t}_{SO}. The lamp must be operated at an ambient temperature of 25°C. The temperature of our laboratory is 23±0.50°C. Therefore, the difference in temperatures is 2±0.50 °C. The difference will cause the standard to decrease by 0.16 ±0.04 °C. If the standard brightness temperature is not corrected by 0.16 °C the total error will not exceed ±0.20 °C. Thus if E^{t}_{SO} is the systematic error caused by the temperature difference, then

$$L^{t}_{SO} = 0.20 \text{ °C.} \qquad (33)$$

Total Source Error, E^{S}_{SO}. The limit on the total source error E^{S}_{SO} in the standard, computed by quadrature addition of (31), (32), and (33), is

$$L^{S}_{SO} = \pm \sqrt{(0.265)^2 + (0.50)^2 + (0.20)^2} \doteq 0.60 \text{ °C.} \qquad (34)$$

Calibration Error, E^{S}_{C}. The calibration uncertainty for the lamp at 1000 °C is given on the certificate as ± 3 °C. Therefore, if L^{S}_{C} is the limit on the standard calibration error,

$$L^{S}_{C} = 3 \text{ °C.} \qquad (35)$$

5.5. Systematic Error in the Test

The following is an analysis of the optical pyrometer for its systematic error.

Alinement Error, E^{T}_{AL}. The brightness temperature of the lamp as measured by the optical pyrometer depends upon p the vertical displacement of the pyrometer with respect to a notch in the filament of the lamp, θ the angle between the normal to the surface of the filament and the optical line of sight, and d the distance between the lamp and the optical pyrometer. Any systematic deviation in p, θ, or d will result in a change in the measured brightness temperature.

Because an increase in d has an opposite effect upon the brightness temperature than does a decrease, a realinement of the pyrometer after each observation will randomly distribute the error caused by the deviation in d among the observations. An increase or decrease in θ or p will cause the observed brightness temperature to increase. Thus, errors caused by deviations in θ or p will be systematic.

It can be shown that p may be alined correctly to within 0.5mm and that θ may be alined to within 0.04 radians. The change in brightness temperature corresponding to a change in θ of 0.04 radians is less than 0.01 °C and the change corresponding to a change in p of 0.5mm is less than 2.00 °C. Thus,

$$L^{T}_{AL} = 2.00 \text{ °C} + 0.01 \text{ °C} = 2.01 \text{ °C} \qquad (36)$$

(Numbers will be rounded to a more realistic number of figures at the end of the calculations.) Interference Error, E^{T}_{IN}. The heat generated by the lamp will cause the temperature of the filter in the optical pyrometer to rise. This will cause the effective wavelength of the pyrometer to change which will result in an error. It has been experimentally determined that the temperature change in the filter is no more than 1 °C. The error caused by such heating is negligible in comparison with the alinement error. Thus,

$$L^{T}_{IN} = 0. \qquad (37)$$

Observer Error, E^{T}_{D}. Because of the relative luminosity curve of one observer usually differs from that of another, there is a systematic difference between measurement results depending on the observer. The limit to the observer error L^{T}_{D} has been evaluated as explained on page 5. It has been estimated to be

$$L^{T}_{D} = 0.5 \text{ °C.} \qquad (38)$$

Sensor Error, E^{T}_{SN}. The sensor error is negligible, therefore,

$$L^{T}_{SN} = 0. \qquad (39)$$

Transmission Error, E^{T}_{T}. Transmission error is negligible because all optical parts are cleaned before a measurement is made, thus

$$L^{T}_{T} = 0. \qquad (40)$$

Indicator Uncertainty, L^{T}_{I}. The least count of the pyrometer used in 10 °F. The readout is in the form of a scale. The readability is taken to be 2.25 °F and the resolution 1.00 °F. Thus, L^{T}_{D} the indicator limit of error is given by

$$L^{T}_{I} = \frac{RD}{2} + RES = 2.13 \text{ °F} = 1.18 \text{ °C.} \qquad (41)$$

Limit to the Total Systematic Error L_s. We obtain the limit L_s to the total systematic error by quadrature addition of (34), (35), (36), (37), (38), (39), (40), and (41) as in (7):

$$L_s = 3.88 \text{ °C} = 6.98 \text{ °F} \qquad (42)$$

5.6. Calibration Data (see table 1)

The procedure previously outlined was used for taking the data to be used in the calibration of the test pyrometer.

Each time step (2) of the procedure was completed, the lamp current was adjusted to give an indicated voltage of 0.053900 as read by the potentiometer. This voltage corresponds to a brightness temperature of 1000 °C. Thus,

$$X^{S}_{i1} = 1000 \text{ °C} = 1832.0 \text{ °F} \qquad (43)$$

for any i, where i is an index which identifies the measurement in the order that it was taken.

The measurement of the standard with the test pyrometer (see step (3)) was taken in the following manner:

One measurement of the brightness temperature of the standard was made by decreasing the current through the reference lamp within the optical pyrometer until the brightness of the reference lamp's filament matched the brightness of the standard. The value obtained from the scale of the test pyrometer was recorded in column 1 of table 1. The other measurement was made by increasing the reference lamp current until a brightness match was obtained. The value of this measurement was recorded in column 2 of table 1.

Column 3 of table 1 is the average of columns 1 and 2 and for any i represents X_{i1}^T the value of the standard as measured by the test pyrometer.

5.7. Computations

The data from column 3 of table 1 was entered into column 1 of table 2.

TABLE 2: Computations

i	X_{i1}^T	$C_{C_i}^T$	$C_{C_i}^T - C_C^T$	$\left(C_{C_i}^T - C_C^T\right)^2$
1	1839.5	- 7.50	-0.50	0.25
2	1839.0	- 7.00	.00	.00
3	1839.0	- 7.00	.00	.00
4	1838.5	- 6.50	.50	.25
5	1839.0	- 7.00	.00	.00
Total	9195.0	-35.00	0.00	0.50
Column	1	2	3	4

Total Calibration Correction Term, C_C^T. The correction term $C_{C_i}^T$ indicated by the ith measurement alone is computed per (11):

$$C_{C_i}^T = \left(X_{i1}^S - X_{i1}^T\right) = 1832.0 - X_{i1}^T \tag{44}$$

The values obtained from eq (44) are entered into column 2 of table 2.

The correction C_C^T is next computed for n, the total number of observations, equal to 5. From eq (10), we have

$$C_c^T = \frac{\text{Sum of col. 2}}{5} - 7.0 \text{ °F.} \tag{45}$$

Limit of the Calibration Error, L_C^T. The deviation of the term C_C^T from $C_{C_i}^T$, $\left(C_{C_i}^T - C_C^T\right)$, is next computed and entered into column 3. The data in column 3 will always sum to zero. The data in Column 3 is squared and entered in column 4. The sum of column 4 is the quantity,

$$\sum_{i=1}^{n}\left(C_{C_i}^T - C_C^T\right)^2$$

which appears in eq (13). When this quantity is multiplied by t^2 and divided by $n(n-1)$, the result will give the square of the calibration uncertainty caused by the random error. For 4 degrees of freedom and a confidence level of 99 percent[2], $t = 4.604$. Therefore

$$\frac{t^2}{n(n-1)}\sum_{i=1}^{5}\left(C_{C_i}^T - C_C^T\right)^2 \tag{46}$$

$$= \frac{21.197}{20}(0.50)\,(\text{°F})^2 = 0.53\,(\text{°F})^2$$

From eqs (13), (42), and (46) we have for L_C^T, the limit of the calibration error for the test pyrometer,

$$L_C^T = (6.98)^2 + 0.53 = 7.02 \text{ °F.} \tag{47}$$

Rounding off to the nearest degree, we have

$$L_C^T = 7 \text{ °F} \tag{48}$$

5.8. Certification

The calibration correction term C_C^T when added to $\bar{X}_1^T$ will agree with $\bar{X}_1^S$ to within $\pm L_C^T$ the calibration uncertainty. From eq (14) we have

$$\bar{X}_1^T = \frac{9195.0 \text{ °F}}{5} = 1839.0 \text{ °F} \tag{49}$$

and

$$\bar{X}_1^S = \frac{5(1832 \text{ °F})}{5} = 1832 \text{ °F.} \tag{50}$$

Substitution of eqs (45), (48), (49), and (50) into eq (12) yields

$$1839.0 \text{ °F} - 7.0 \text{ °F} = 1832.0 \text{ °F} \pm 7 \text{ °F.}$$

Thus it is seen that when -7.0 °F is added to the scale of the optical pyrometer the resulting value will be correct to within ±7 °F.

6. Summary

After the presentation of essential concepts and formulas it has been shown how to combine individual error limits to obtain the limit of the total error in a calibration.

The author is indebted to a referee for the correction of many errors, obscurities, and phrasings.

7. References

1. P. G. Hoel, Introduction to Mathematical Statistics, John Wiley, & Sons, Inc. New York, 1947.

2. R. A. Fisher and F. Yates, Statistical Tables for Biological, Agricultural and Medical Research, Table III, Hafner Publishing Co., Inc., New York, 4th ed., 1953.

3. C. R. Barber, factors affecting the reproducibility of brightness of tungsten strip lamps for pyrometer standardization, J. Sci. Instr. 23, 238 (1946).

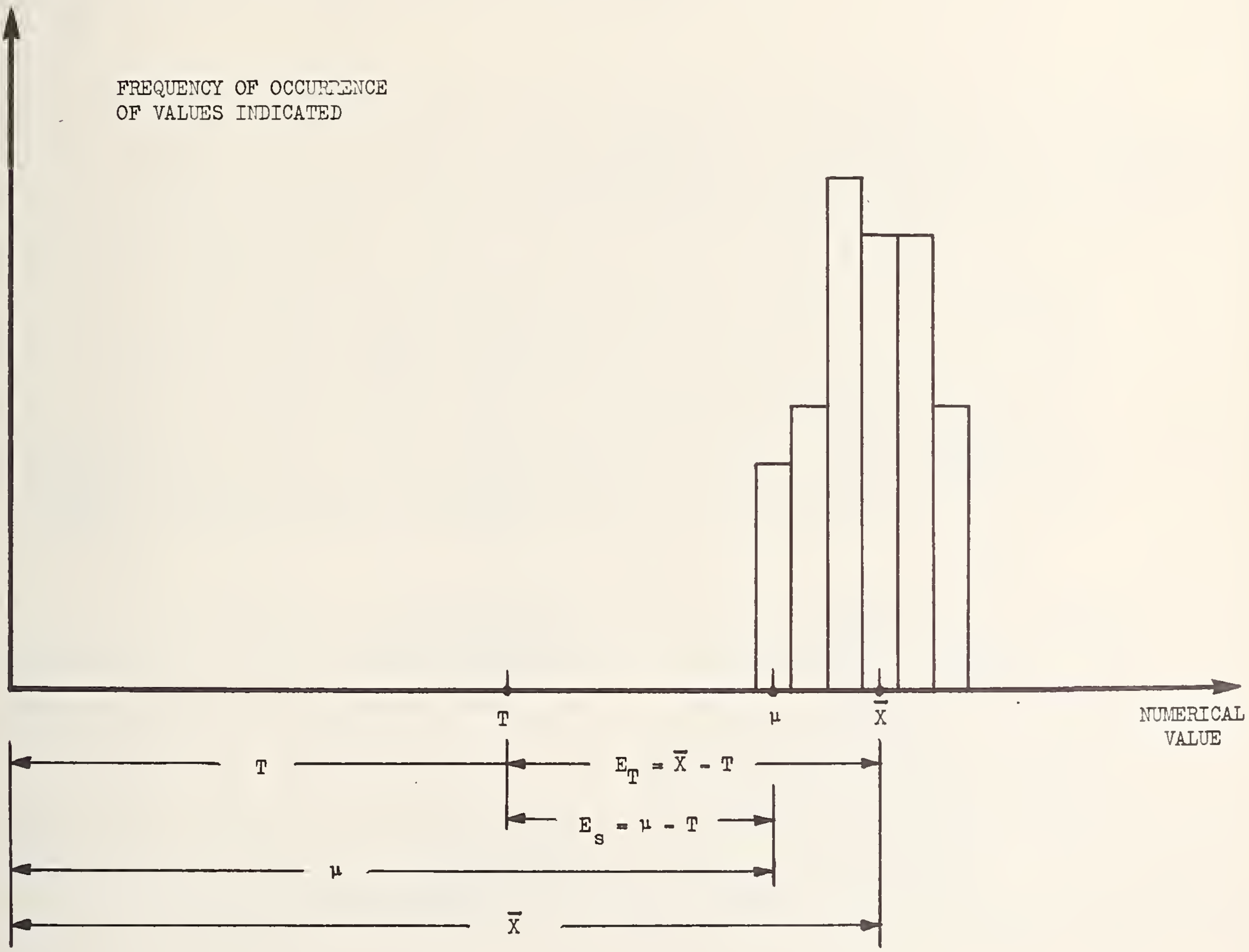


Figure 1 - SCALE OF VALUES

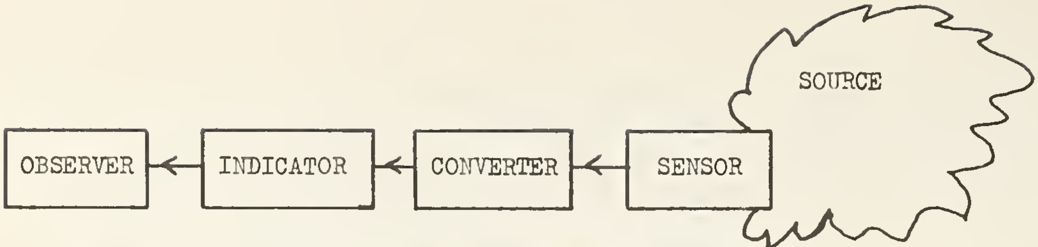


Figure 2 - "The Generalized Measurement System"

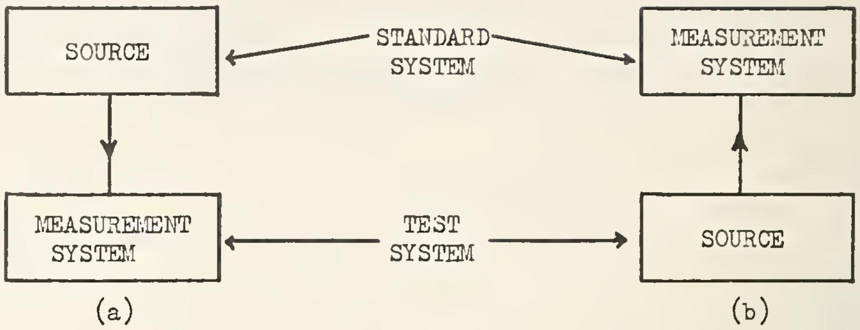


Figure 3

Figure 4

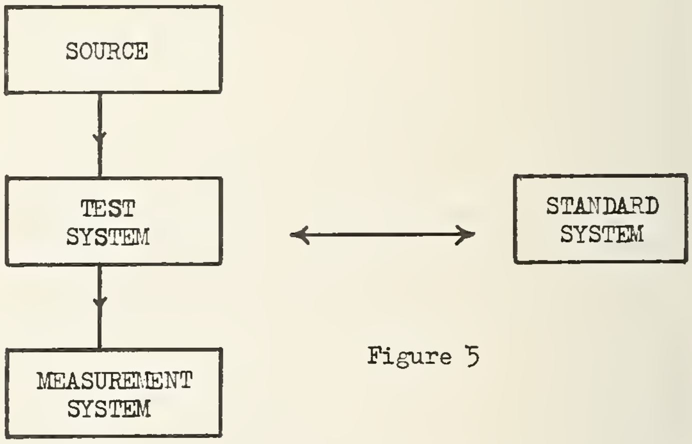


Figure 5

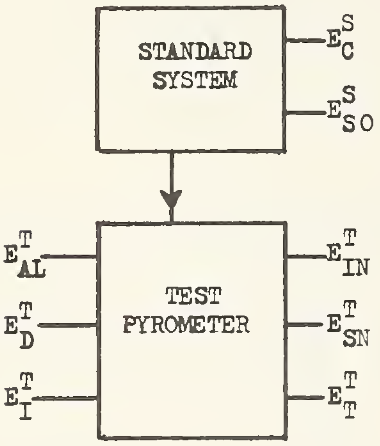


Figure 6 - "Calibration System"

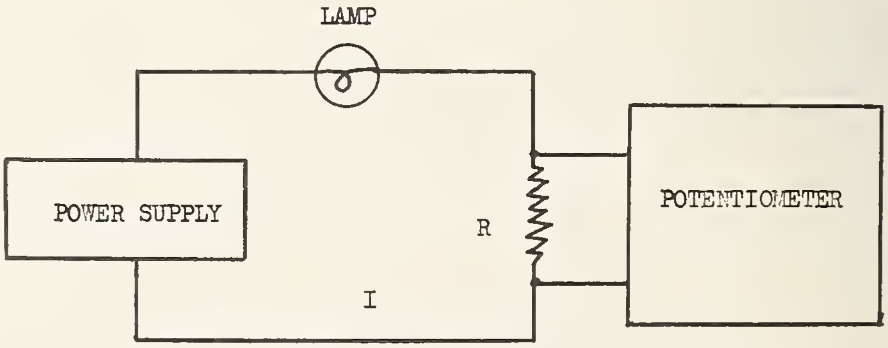


Figure 7 - "The Standard System"

NBS Miscellaneous Publicotion 248
Proceedings of the 1962 Standards Laboratory Conference

Session 2. Error Analysis of Measurement Systems

Paper 2.4. Measurement Errors—Identification, Detection, Evaluation and Expression

Arthur J. Plourde *

In an effort to achieve compatible statements concerning measurement uncertainty within the Navy Calibration Program, the Metrology Department of the Bureau of Naval Weapons Representative, Pomona, California, is presently publishing a document which will define concepts of measurement uncertainty and recommend methods for the evaluation of the uncertainty. The following text is a summary of that document, which will appear as an Engineering Circular.

Procedures, techniques and other measurement guides, environmentally controlled laboratories, high quality equipment and well trained measurement technicians all contribute to make measurements of high precision possible. However, accuracy, which depends on the unique conditions at the time of the measurement and on the metrologist observing them, cannot easily be, and has not been, controlled or even well defined by a single responsible agency.

The major problem in error analysis is the systematic error. The statistical treatment of data allows everyone to express the random errors in his measurements in a common language. In the case of the systematic error, however, it is nearly impossible to find two persons who will give the same sources and corresponding magnitudes of errors in a given measurement.

1. Introduction

The most important single item of a measurement is the report of the results. The report is the product the customer is paying for. The report is what the customer must use as a foundation for his work, whatever it may be. The report must, therefore, be valid, meaningful, and unambiguous.

The validity depends upon the measurement technique used, the quality of the measurement system, and the ability of the observer. The report is meaningful only when the results are given for conditions of actual use of the test item and when the uncertainty of the result due to errors is included. An unambiguous report requires that the results and the expression of the uncertainty be given in simple and well-defined terms.

This text is concerned solely with the expression of the uncertainty in the reported value.

Metrologists everywhere are appealing for universally acceptable concepts of the total measurement uncertainty; concepts which will result in the recommendation of a single method for estimating and expressing the measurement uncertainty. What has prevented even a proposal of such concepts and recommended methods? The answer can be found in the editorial in the February 1961 issue of the ISA Journal, which was devoted to Measurement

---

*Vitro Corporation of America, Pomona, Calif.

Standards or, perhaps more appropriate, the measurement gap. In that editorial, Mr. Covey made an appeal for seeking, encouraging, and releasing 'new concepts of government, education, communication, and transportation . . . .' He concludes his editorial by saying, "Perhaps the biggest problem will not be to generate new concepts, but to overcome the pressures and resistances of tradition, habit, dictatorial management, politics, and complacency."

2. Classification of Errors

The uncertainty of a measurement is the interval about the quoted value in which the mean, actual, or true value of the unknown is believed to lie. An uncertainty is associated with the result because one can never exactly determine measurement error. The uncertainty is thus the interval of a probable range of the error in the measurement. The probability is equal to the confidence associated with the interval. Universal concepts of errors are, therefore, basic to universal concepts of uncertainty. In the Navy Calibration Program errors are classed according to their effect on the result [Stout, 1960].

2.1. Gross Errors

Errors which completely invalidate or drastically reduce the reliability of the result are classed as gross errors. Such things as outright mistakes, calculation blunders, misreading of scales, equipment failure, and invalid procedures are causes of gross errors. This class of error is never reported, for when gross errors are noted the corresponding readings are omitted from the data or the measurement is repeated. The rejection of large errors (actually large deviations from the mean) in the original data by use of statistical tests is based on the assumption that they are caused by gross errors.

2.2. Systematic Errors

Errors which tend to bias the result of a measurement are classed as systematic errors. Factors which cause systematic errors include the ambient environment, the calibration of the measurement equipment, the measurement technique employed, and the observer. In most cases, there is a known mathematical relationship between the unknown being measured and the influencing factors. When the directions and magnitudes of the systematic errors can be determined, they are removed from the data. No mention is made in the calibration report concerning such corrections. When the directions and magnitudes cannot be determined, a value of the uncertainty due to possible systematic errors must be estimated and included in the report. The estimation of the systematic error is discussed in section 3.

2.3. Random Errors

Errors which cause differences in repeated values of the same measurement conducted under controlled conditions are classed as random errors. The cause of the random errors is the uncontrollable and unnoticed fluctuation of the many influencing factors of the measurement [Eisenhart, 1961]. Random errors may be thought of as high-speed or short-term systematic errors. Since there are a large number of influencing factors, the data for the measurement are approximately normally distributed about the mean [Mood, 1950]. Hence, the random error in a measurement is defined to have a normal probability distribution with zero mean and finite variance[1].

3. Estimating the Uncertainty Due to Systematic Error

The uncertainty due to the systematic error is estimated by use of the classical uncertainty propagation formula for those factors which have a known mathematical relationship to the unknown variable [Beers, 1958; Frank, 1959; Stout, 1960]. The value assigned as the uncertainty of each factor should be based on an analysis of the system measuring and/or controlling that factor. Thus, if temperature is one of the influencing factors in a measurement and the control system is a regulated oil bath, the metrologist must determine the possible variations in temperature due to cycling of heaters, gradients in the oil, and inaccuracies of the indicating device. The degree to which each factor must be analyzed depends upon the variation of the factor and its relationship to the unknown variable.

The uncertainty due to systematic error for factors which have no known mathematical relationship to the unknown variable (such as observer bias) must be determined from experience; i.e., analysis of a large number of measurements. The magnitude assigned to such factors is added to the systematic error uncertainty calculated from the uncertainty propagation formula. This conclusion is based on an analysis of the measurement represented by the mathematical model

$$Yi = \mu_y + \sum_{j=1}^{k} \left(\delta y / \delta x_j \right)_0 dx_j + \Delta , \qquad (1)$$

where yi = measured value of the variable,
μ_y = actual value of the variable,
$\left(\delta y / \delta x_j \right)_0$ = partial of y with respect to X_j evaluated at the indicated (measured) values of the x's,
δx_j = error in X_j (limiting values),
and Δ = systematic error in y due to factors without a known mathematical relationship to y.

Defining the systematic error as the root of the expected value of the square deviations of the

---

[1] Variance is a measure of variability used by statisticians; Metrologists use the standard deviation which is the square root of the variance.

measurements for factors whose magnitudes and direction are not exactly known, it follows from (1) that

$$U_s = \left[\sum_{j=1}^{k} \left({}^{\delta y}/{\delta x_j} \right)_0^2 \; dx_j{}^2 \right]^{\frac{1}{2}} + \Delta \tag{2}$$

where U_s systematic error uncertainty and all other terms are as previously defined. Since the last term in (2) requires a good deal of analysis on a measurement before its magnitude can be determined, it is being omitted (but not forgotten) in the present error analysis scheme in the Navy Calibration Program. The uncertainty due to factors with no known mathematical relationship to the unknown variable will show up as between measurement effects. Systematic error uncertainty is determined for the present by use of the classical uncertainty propagation formula,

$$U_s = \left[\sum_{j=1}^{k} \left({}^{\delta y}/{\delta x_j} \right)_0^2 \; dx_j{}^2 \right]^{\frac{1}{2}} \tag{3}$$

It is often stated, without qualification, that uncertainties combine by quadrature addition. This statement is often taken as a rule rather than a result. Equation (3) results in quadrature addition of proportional uncertainties for some of the more common expressions. Consider, for example, a variable which is related to the various factors by their product,

$$Y = f(x's) = \prod_{j=1}^{k} X_j \tag{4}$$

The uncertainty of y as computed by (3) is

$$U_s \left[\sum_{j=1}^{k} \left(\prod_{i \neq j}^{k} X_j{}^2 \right) dx_i{}^2 \right]^{\frac{1}{2}} \tag{5}$$

Expressing the error of the $x's$ in proportional parts, (5) becomes

$$U_s = \left[\sum_{i=1}^{k} \left(\prod_{j=1}^{k} X_j{}^2 \right) \left(\frac{dx_i}{x_i} \right)^2 \right]^{1/2} \tag{6}$$

By factoring and simplifying, (6) can be written as

$$U_s = y \left[\sum_{j=1}^{k} Ux_j{}^2 \right]^{1/2}.$$

or (7)

$$U_{s/y} = \left[\sum_{j=1}^{k} Ux_j{}^2 \right]^{1/2},$$

where Ux_j = proportional part uncertainty of x_j.

Quadrature addition of uncertainties here is obviously a result and not the basic rule. Many calibrations are a comparison of an unknown to a standard. In this case the mathematical relationship is

$$y = S f(X'S), \tag{8}$$

where S = value of standard
and $f(X'S)$ = comparison technique.

Performing the same operations on (8) as were performed on (4), the uncertainty of (8) is given as

$$U_s = y \left[U_c^2 + U_{f(x's)}^2 \right]^{1/2}. \tag{9}$$

where U_c = proportional part uncertainty of the certified value of the standard,
and $U_{f(x's)}$ = proportional part uncertainty of the comparison technique used.

Here again quadrature addition is a result and not the rule. It can be demonstrated that simple quadrature addition of uncertainties is not a general result [Ramboz, 1961]. An example of the application of the foregoing is given in appendix A.

Detection of large errors which may be reduced will be a by-product of the analysis of the systematic error uncertainty,

4. Estimating the Uncertainty Due to Random Error

The classical uncertainty propagation formula has been shown to be excellent for determining the uncertainty in a result which is a function of a number of random variables, especially when the partial effects of the variables are linear [Tukey, 1958]. For independent errors the propagation formula is

$$\sigma_y^2 = \sum_{i=1}^{k} \left({}^{\sigma y}/{\sigma x_i} \right)_0^2 \; \sigma_{x_i}^2, \tag{10}$$

where $Y = f(X'S)$,
σ_x^2 = variance of X,
$\left({}^{\delta y}/{\delta x_i} \right)_0$ = partial of y with respect to x_i; evaluated at the average values of the $x's$.

While (10) yields excellent results, its use in the laboratory is not as practical as using the measurement data of the result. The random error in the result can be estimated from the measurement data. An uncertainty due to the random error can be assigned if one assumes a given probability distribution for the result. Since measurement variables satisfy the conditions of the Central Limit Theorem[2] and the calibration value is the mean of a set of readings, the probability distribution of the result is assumed to be normal with a finite mean and variance to be determined from the measurement data.

The calibration value is not always a mean but rather some adjusted value of the measurement data. For repeated readings of one value, the least squares adjustment is the mean. This is the most widely used adjustment. However, higher order types of adjustment are often needed as in the case of the calibration of flowmeters and attenuators. The Metrology Department feels that adjustments of data beyond that of a straight line should be minimized until the laboratory personnel are educated in the basic concepts of statistics.

As stated previously, the measured values of the unknown variable are considered to be normally distributed and the calibration value given as the result is the mean of the measured values. The uncertainty of the result due to random error is defined by the Metrology Department to be the 95 percent confidence interval of the mean. Evaluating this interval requires the use of the standard deviation of the sample (measurement), the sample size (number of readings), and the Student -'t' distribution. This same information must be available to the customer for a correct understanding of the quoted uncertainty.

The mathematical expression for the 95 percent uncertainty interval of the mean due to random error is

$$P\left[\overline{X} + S_{\overline{x}}\,(n-1)t\,0.025 \le \mu \le \overline{X} \right. \qquad 11$$

$$\left. + S_{\overline{x}\,(n-1)t}\,0.975 \right] = 0.95,$$

where $\overline{X}$ = mean of the measurement data,
$S_{\overline{x}}$ = standard deviation of the mean,
$(n-1)t\,0.975$ = tabulated value of the one-tail Student- 't' distribution for (n-1) degrees-of-freedom at a confidence level of 0.975,
n = number of readings,
and 0.95 confidence level.

This expression reads: ''the probability that the true value lies between the mean, $\bar{x}$, plus the product of the standard deviation of the mean, $S\overline{x}$, times the 't' value for $(n-1)$ degrees-of-freedom at a confidence level of 0.025 and $\bar{x}$ plus the product of the standard deviation of the mean times the 't' value for $(n-1)$ degrees-of-freedom at a confidence level of 0.975 is 0.95.'' An example of the foregoing applied to actual measurement data is given in appendix B.

5. Expressing the Measurement Uncertainty

The choice of a 95 percent confidence level for the uncertainty due to the random error was not completely arbitrary. The choice was made with consideration to the fact that the random error uncertainty and the systematic error uncertainty must be combined to give the total measurement uncertainty. This combination must be with uncertainties at the same level.

Usually the uncertainty due to the systematic error is given at a 100 percent confidence level. This is done by multiplying a calculated systematic uncertainty by a factor of 3 or 4. Such a practice results in the inflation of the uncertainty interval until the end item cannot be calibrated to the required accuracies through the necessary echelons of standards laboratories. Since the calculated systematic error uncertainty is usually increased, it must be assumed that it is not at the 100 percent confidence level. The question is then, what confidence can be assigned to the calculated systematic error? To determine this formally may be impossible; consider, however, a somewhat logical analysis as follows:

(a) As the uncertainty interval is increased, the confidence level approaches 100 percent.
(b) Since there are normally a large number of techniques for making the measurement, the selection of one indicates there is a high confidence associated with it.
(c) The magnitudes of the factors used in calculating the systematic error are limiting values.

Based on these considerations, a 95 percent confidence for the calculated systematic error does not seem unreasonable. Certainly it is better to assume this confidence than to needlessly inflate the uncertainty statement.

Total measurement uncertainty in the Navy Calibration Program is expressed (and defined) as the quadrature addition of the calculated systematic error (calculated by use of (3) and the 95 percent confidence interval for the mean value of the measurement data (calculated by (11)). The above statement is based on a model of a measured value given by

$$y_i = \mu_y + \sum_{j=1}^{k} \left({\delta y}/{\delta x_j} \right)_0 dx_j + \Delta + r_i\,, \qquad (12)$$

where r_i = random component in y_i and all other terms are as defined in (1).

The expected value of the square-deviations of (12) is

$$E(SD) = \left[\left(\sum_{j=1}^{k} \left({\delta y}/{\delta x_j} \right)_0^2 dx_j^2 \right)^{1/2} + \Delta \right]^2 + \sigma_y^2\ . (13)$$

---

[2] The Central Limit Theorem states that ''if an arbitrary population distribution has mean μ and finite variance σ^2, then the distribution of the sample mean approaches the normal distribution with mean μ and variance $\sigma^2 n$ as the sample n increases.''

Replacing the terms in (13) with the 95 percent confidence uncertainty intervals, the uncertainty 95 percent confidence) of a measurement is found to be

$$U_y = \left[U_s^2 + U_r^{\,2} \right]^{1/2} . \qquad 14$$

where U_y = measurement uncertainty (95% confidence),
U_s = systematic error uncertainty,
and U_r = random error uncertainty.

An example expressing the total measurement uncertainty is given in appendix B.

6. Appendixes

Appendix A. Systematic Error Analysis of a Mass Certification System

The mass of an unknown item is to be determined by comparison to a standard mass on an equal arm balance. The procedure is to balance (by computing the rest points) both the standard and the unknown against a tare weight. The rest point is determined by the formula

$$R = \frac{x_3 + 1/2\,(x_2 + x_4)}{2},$$

where R = rest point, (A. 1)
x_1 = scale reading at end point of swing (x_i reading omitted in all cases),

and the scale is numbered consecutively from right to left.

In most cases, the rest point for the standard and unknown will not be the same; therefore, the sensitivity of the balance must be determined. This is done by adding a small mass to the heavier pan and calculating the sensitivity by the formula

$$\text{Sen.} = \frac{s_3}{(R_2 - R_1)} \qquad (A.2)$$

where Sen. = sensitivity,
S_3 = added mass,
R_1 = rest point for unknown without added mass,
R_2 = rest point for unknown with added mass and no change is made in carrier and beam weights.

The unknown mass is determined by the equation

$$Y = S \left[1 + \frac{C_2 - C_1}{S} + \left(B_2 - B_1\right) \frac{S_2}{S} + \left(\frac{R_3 - R_1}{R_2 - R_1}\right) \frac{S_3}{S} \right] \qquad (A.3)$$

where Y = unknown mass,
S = standard mass,
B_1 = relative position of beam mass for unknown,
B_2 = relative position of beam mass for standard,
S_2 = value of beam mass,
C_1 = carrier mass for unknown,
C_2 = carrier mass for standard,
R_3 = rest point for standard, and other terms are as in (A.2).

Equation (A.3) is a form of the comparison equation given in (8).

The uncertainty of (A.3) due to systematic error as determined by use (3) is

$$U_s = \left[ds^2 + 2\,dc^2 + \left\{ \left(B_2 - B_1\right)^2 ds_2^{\,2} + 2S_2^{\,2}\,dB^2 \right\} + 2 \left\{ \frac{s_3}{(R_2 - R.)} \right\}^2 dR^2 + \left(R_3 - R_1\right)^2 \left\{ \frac{ds_3^{\,2}}{(R_2 - R_1)^2} + \frac{s_3^{\,2}\,dR^2}{(R_2 - R_1)^4} \right\} \right]^{1/2} \qquad (A.4)$$

where ds = uncertainty of standard mass,
dc = uncertainty of carrier mass,
ds_2 = uncertainty of beam mass,
dB = error in relative beam mass position,
dR = uncertainty in rest point,
and ds_3 = uncertainty of sensitivity mass.

The term/ $\left\{(B_2 - B_1)\, dS_2^{\,2} + 2S_2 dB^2\right\}^{1/2}$ is the uncertainty due to the beam mass and position. The term $\left\{ \frac{dS_3^{\,2}}{(R_2 - R_1)} + \frac{S_3^{\,2}\, d\mathrm{R}^3}{(R_2 - R_1)^4} \right\}^{1/2}$ is the uncertainty in determining the balance sensitivity. The coefficient 2 in some terms on the right-hand side of (A.4) result from the fact that these uncertainties enter the calculation twice, once for the unknown and once for the standard. The uncertainty of the rest point determination as determined by application of (3) to (A.1) is

$$dR = \left[1/2\ d_x^{\,2} \right]^{1/2} \qquad (A.5)$$

where d_x is the uncertainty of reading the end point swing position.

Consider the systematic error in the mass determination of a 100 gram item. In this case, values are listed below:

ds = 0.0005g
dc = 0.00003
ds_2 = 0.00005g
ds_3 = 0.00011
S_2 = 0.5g
R_2 = 14.25
B_1 = 0.1

dx = 0.1
dR = 0.07
dB = 1 x 10^{-5}
S_3 = 0.001g
R_1 = 20.1
R_3 = 17.3
B_2 = 0.9

The uncertainty as computed from (A. 4) is

$$U_s = \left[25 \text{ x } 10^{-8} + (2)(9) \text{ x } 10^{-10} + \left\{ (64)(25) \text{ x } 10^{-12} + (2)(25) \text{ x } 10^{-12} \right\} + 2(3.2) \text{ x } 10^{-12} + (9.3) \left\{ (3.2 \text{ x } 10^{-12} + (.8) \text{ x } 10^{-12} \right\} \right]^{1/2} \qquad (A.6)$$

$$= \left[25.41 \text{ x } 10^{-8} \right]^{1/2}$$

$$= 5.1 \text{ x } 10^{-4} \text{ grams.}$$

Appendix B. An Example of Estimating Uncertainty Due to Random Error and Expressing the Total Uncertainty

Table 1. Measured values of 100-g tare mass

99.7833	99.7833
99.7830	99.7834
99.7834	99.7835
99.7833	99.7832
99.7835	99.7833

A set of ten measurements of a 100-tare mass was made using the procedure given in appendix A. The results are given in table 1.

Computations for the mean, standard deviation of individual values, and standard deviation of the mean are given in table 2.

The 95 percent confidence interval for the mean of these data as computed by (11) is 1.1 x 10^{-4} grams. In this example the value of t is 2.262 (see table 3).

The product of t and $S_{\bar{x}}$ is 1.1 x 10^{-4}. Placing this value and the value of $\bar{X}$ in (11), the probability statement for the mean of these data is

$$P\left[99.78321 \leq \mu \leq 99.78343\right] = 0.95 . \qquad (B.1)$$

Table 2. Computations

i	$X_i \times 10^{-4}$ *	$(X_i - X) \times 10^{-4}$	$(X_i - \bar{X})^2 \times 10^{-8}$	Calculations
1	33	-0.2	0.04	$\bar{X} = \frac{332}{10} \times 10^{-4} = 33.2 \times 10^{-4}$
2	30	-3.2	10.24	
3	34	+0.8	0.64	$S_x = \left[\frac{19.60}{9} \times 10^{-8}\right]^{1/2} = 1.47 \times 10^{-4}$
4	33	-.2	.04	
5	35	+1.8	3.24	$S_{\bar{x}} = \frac{1.47 \times 10^{-4}}{10} = 0.47 \times 10^{-4}$
6	33	-0.2	0.04	
7	34	+.8	.64	
8	35	+1.8	3.24	
9	32	-1.2	1.44	
10	33	-0.2	0.04	
Sum	332×10^{-4}	0**	19.60×10^{-8}	

*These are coded values. (X_i + 99.78 = actual value.)
**This sum should always equal zero.

The total uncertainty of the measurement of mass as calculated by (12) is

$$U = [(5.1 \times 10^{-4})^2 + (1.1 \times 10^{-4})^2]^{1/2} \qquad (B.2)$$
$$= [27.22 \times 10^{-8}]^{1/2}$$
$$= 5.2 \times 10^{-4} \text{ grams.}$$

In the Navy Calibration Program the recommended manner for stating the foregoing findings is to identify the activity requesting the measurement, the activity issuing the report, and the test item. The mean value (99.7833) is given as the value of the unknown followed by the statement, "the above value is the mean value of a set of ten measurements. The total uncertainty of this value, including random error uncertainty, is 5.2 x 10^{-4} grams (5.2 ppm)."

Table 3. Some Values of t

(Probability/t/ < γ = 0.05)	
Degrees of freedom (n - 1)	γ
7	2.365
8	2.306
9	2.262
10	2.228

7. References

Beers, Yardley, Introduction to the Theory of Errors, 2d ed., Addison-Wesley Publishing Co., Inc., 1958.

Eisenhart, Churchill, Expression of the uncertainties of calibrations, working paper of the Subcommittee on Accuracy of NBS Testing, September 1961.

Frank, Ernest, Electrical Measurement Analysis, McGraw-Hill Book Co., Inc., 1959.

Mood, A., Introduction to the Theory of Statistics, 1st ed., McGraw-Hill Book Co., Inc., 1952.

Ramboz, John D., A Statistical method for determining measurement system uncertainties, paper presented at ISA Conference and Exhibit, September 1961.

Stout, Melville E., Basic Electrical Measurements, 2d ed., Prentice-Hall, Inc., 1960.

Tukey, John W., The propagation of errors, fluctuations, and tolerances—basic generalized formulas, Tech. Report No. 10, Princeton University, Dept. of Army Project 5B-99-01-004, Contract No. DA 36-034-ORD 2297. Estimated date of report, 1958.

Session 3. NCSL Business and Information

Annual Report of the Chairman of the National Conference of Standards Laboratories

L. B. Wilson *

1. Introduction

The advance program for this conference, in the section describing Session 3, contained a statement which read as follows:

> "During the past year we have been proceeding along parallel courses to make NCSL a going organization. One course is to get organization, membership, liaison, and sponsorship details of NCSL established. The other course is to get action items going to meet some of the needs of standards laboratories."

This session was originally intended to give details about each of these parallel courses, supplementing them with a briefing on the results of the General Committee meeting which was held just prior to the conference. However, in the period of time since this program was prepared, we worked out the details of a Delegates' Assembly. The Delegates' Assembly is going to cover certain business items which had originally been planned for this session. You will still get a general description of NCSL's business in this session and, if you attend the Delegates' Assembly, you will get it from a slightly different point of view, with some parts being given in more detail.

2. Plans for the Delegates' Assembly

You are all welcome to attend the Delegates' Assembly tomorrow afternoon. It will be held in this auditorium. Those of you who have been officially certified as delegates will sit in the front in a separate area and will be entitled to vote. The rest of you can attend as observers and sit towards the rear of the auditorium.

If your company didn't receive a notice of this Delegates' Assembly, or if for any reason there is no official delegate from your company, this can still be taken care of. There are forms at the main desk in the lobby which can be filled out to certify a delegate, even if it means self-certification in some cases. We do require that there be not more than one delegate from each company. When we speak of "company" we mean a major plant at one geographical location. If you have plants in several different parts of the country, there can be one delegate from each plant. However, we don't want to have two delegates from any one plant at one street address. If there is any question on this, see me or see Harvey Lance, and I think we can work the details out.

One of the major items of business of the Delegates' Assembly will be the election of officers and General Committee members for the next year. In addition, I will briefly describe our activities of the past year and make some suggestions for the coming year's activities.

3. NCSL Bylaws

I want to call your attention to the fact that the NCSL does have bylaws. Copies were given out with the packet of reprints which you received when you registered. If you didn't get one, there are additional copies at the main desk in the lobby.

4. NCSL Participation and Membership

Some of you have asked how you can participate in NCSL work or how you can become members. So far we do not have a formal type of membership in NCSL. In other words, you cannot apply for membership in NCSL and be accepted, either on an individual basis or a company basis. This is still one of the parallel course organizational details which we are developing. I think the

---

*Sperry Gyroscope Company, Great Neck, New York.

important thing here is that we are going along these parallel courses. We have gotten ourselves organized to the point where we do have a set of bylaws, six pages long. In addition we are working on some action items which are significant. In another six months or a year, we hope to have some of the further organizational details, such as membership, worked out.

In the meantime, "membership" is pretty much on the basis of your participation in NCSL activities, such as attendance at this conference, attendance at some of our Standards Laboratory Management Workshops, or by your activities as members of NCSL committees. We had six one-day Standards Laboratory Management Workshop earlier this year, on a trial-run basis, and I anticipate that we will be having more in the coming year. If you are not already on an NCSL committee and want to get on one, you may have a chance after this evening's session. Don't be surprised if some of the committee chairmen sitting at the table here with me say "I'm looking for members for my committee; if you are interested, please see me."

We also have an Interest Survey Questionnaire, copies of which are on the table in the lobby. Please fill one of these out if you are interested. Note that there is a place on the back where you can check the particular subjects which you are interested in. The filled-out questionnaires will be routed to the various NCSL committee chairman as suggestions for additional people for their respective committees. After you have filled out the questionnaire, you can either give it to me or leave it at the main desk in the lobby.

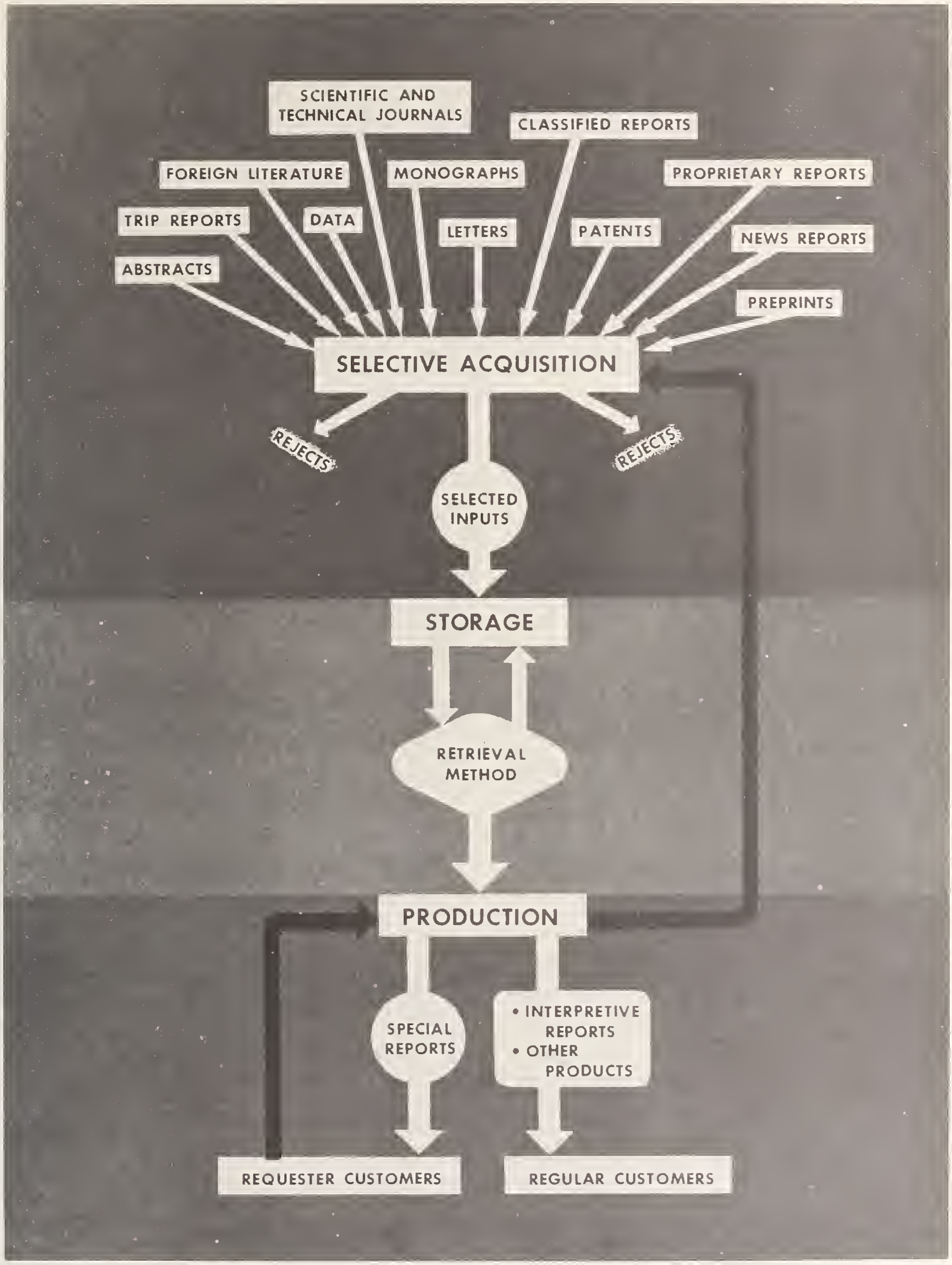
5. The Nature of NCSL's Business

To help us understand the nature of NCSL's business, I will read a portion of our bylaws, from Section II, Purposes and Functions:

> "A. The NCSL is an organization to promote cooperative action on common problems of management and operation of measurement standards and calibration laboratories.
>
> B. The NCSL pursues its goal by all appropriate means through joint, voluntary cooperation of its participants as follows:
>
> 1. Holding conferences, workshops, seminars, and meetings for presentation of papers and discussions pertaining to technical and managerial problems, operating practices, and policies for standards laboratories.
> 2. Collecting and disseminating information about current practices for the organization, operation, and evaluation of measurement standards and calibration laboratories.
> 3. Preparing and disseminating source material and recommendations for preferred technical procedures and operating practices.
> 4. Providing a medium of interchange of information relative to deficiencies or advances in calibration techniques and laboratory instrumentation.
> 5. Collecting, analyzing, and disseminating pertinent statistical information related to the scope, growth trends, operation, and management of measurement standards and calibration activities.
> 6. Assisting measurement standards and calibration laboratories in arranging measurement agreement audits."

The above group of six items can be broken down into two major parts. The first of these parts is the No. 6 item--"assisting measurement standards and calibration laboratories in arranging measurement agreement audits." One of our conference sessions tomorrow morning will take up this particular subject; it will be Session 5, Measurement Agreement Comparisons Among Standardizing Laboratories. Therefore, I will touch on this subject briefly at this time by pointing out that there are two basic types of measurement agreement. One of these is the type where NBS does not provide calibration services and where we might want to have a measurement agreement round-robin, or perhaps even a series of round-robins, to establish a certain degree of agreement in measurements. The other type is one where NBS does provide calibration services, and which is used when we want to see how good our measurement agreement is, rather than to establish measurement agreement. In other words, this second type would be an auditing operation to determine how well we are achieving agreement, not only at the standards laboratory or calibration laboratory level, but all the way down to the end-use level--i.e., the measurement or test level. I think that providing assistance for both of these types of measurement agreement is a major part of NCSL's business.

The second major part of NCSL's business is something which I will lump under the general heading of information dissemination. This includes the first five items under Section IIB of our bylaws--i.e., holding conferences, collecting and disseminating information, preparing and disseminating source material, proving a medium of interchange, and collecting, analyzing and disseminating pertinent statistical information. I am not going to discuss these subjects now because they will be covered later in this evening's session. (A copy of the printed Annual Report distributed at the Conference is attached to these Proceedings as Appendix 1 of this session).



The Primary Functions of Scientific Information Centers

(Reprinted with permission from Battelle Memorial Institute).

Session 3. NCSL Business and Information

Appendix I. Annual Report of the Chairman of the General Committee 1961-1962

This report covers the period from September 15, 1961 to August 1, 1962. The date of September 15, 1961 is an important one because it was then that the NCSL was established officially in Los Angeles at a Standards Laboratory Conference arranged by an Ad Hoc Committee for an Association of Standards Laboratories. However, to provide background for many of the things I have to say, I am also going to discuss the series of events prior to this which resulted in the establishment of the National Conference of Standards Laboratories.

Origin of the NCSL

The real origin of the National Conference of Standards Laboratories goes back to the June 1960 Conference on Standards and Electronic Measurements held at the Boulder Laboratories of the National Bureau of Standards. The suggestion of the need for some kind of an association of standards laboratories was brought out here in a paper by Mr. Harvey W. Lance, Chief of the Electronic Calibration Center at NBS Boulder. Entitled "The Nation's Electronic Standards Program: Where Do We Now Stand?", the paper discussed many problems involved in the management of standards laboratories, including such things as what constitutes a good standards laboratory, traceability of calibrations, justification of standards requirements, interim standards and calibration services, measurement agreement, self-qualification of standards laboratories, education and training of standards personnel, and the electronic calibration services of NBS. Mr. Lance ended his paper by suggesting the need for an association of standards laboratories to work with NBS to solve these problems.

The next day, June 23, approximately 125 of the conference attendees met to discuss Mr. Lance's suggestion further. There was unanimous agreement that some action should be taken, and as a result this group requested the Boulder Conference Committee to establish an Ad Hoc Committee. Mr. H. C. Biggs of Sandia Corporation was appointed to head the Ad Hoc Committee which was soon increased to twenty members with representatives from the Department of Defense, the three military services, Atomic Energy Commission, NBS, industrial standards laboratories, instrument manufacturers, and educational institutions. The task of this Ad Hoc Committee was to decide on the need for and the means of establishing an organization of standards laboratories such as Mr. Lance had proposed.

The first meeting of the Ad Hoc Committee was held in New York City on September 27, 1960. This was an exploratory meeting to try to define the problems faced by standards laboratories and to see why these problems could not be solved by one or more of the many organizations already in existence--particularly organizations such as the Instrument Society of America, the Institute of Radio Engineers, the American Institute of Electrical Engineers, or the American Standards Association. The conclusion reached at this meeting was that what was missing was a medium of exchange at the management level to work on the problems of standards laboratories and to establish liaison with technical societies and other organizations for help in solving these problems. This, in turn, seemed to indicate the need for an association of laboratories, not of people, thereby justifying Mr. Lance's original suggestion for formation of an association of standards laboratories.

The second meeting of the Ad Hoc Committee was held in Albuquerque, New Mexico, on February 13 and 14, 1961. At this meeting it was decided that there was a definite need for an organization of standards laboratories, that it should concentrate on developing voluntary formats and procedures, and that it should stress the collection and dissemination of information relative to calibrations and standards. The major objectives of such an organization of standards laboratories should include a continuing study to determine the requirements and needs for calibrations and standards, improvement of measurement abilities to meet these requirements, and establishment of measurement agreement operations to determine how well the requirements are met.

An interesting observation in the minutes of the Albuquerque meeting was the statement that the vast majority of future calibrations at the working level must of necessity be conducted by rather small calibration laboratories and that accuracy at this level must be assured. Although this conclusion is obvious to most of those engaged in precision measurement work, it is well to note that the needs of the smaller laboratories were recognized as early as this during the formative stages of the NCSL.

A decision of the Albuquerque meeting was to streamline the Ad Hoc Committee by the appointment of an Executive Committee. Such an Executive Committee was established, and it met at NBS Boulder on May 23, 1961. It established four working subcommittees and during the summer of 1961 these committees made plans for a one day standards laboratory conference to be held in September, 1961, in conjunction with the ISA Annual Conference in Los Angeles.

During the Standards Laboratory Conference at Los Angeles, on September 15, 1961, each of the four working subcommittees presented reports of their preliminary studies and recommendations for future activities. The purpose of this was not only to take some steps towards solution of the problems of standards laboratories, but also to present samples of the results which might be expected if an organization of standards laboratories were established and if future standards laboratory conferences were held.

The last part of the September 15, 1961 Standards Laboratory Conference was a business meeting which stressed the advantages of a "conference" type of organization as a compromise between a continuation of the present Ad Hoc Committee and a more formal "association" type of organization. Mr. William Wildhack, Associate Director of NBS, stated that NBS had sponsored a number of conference-type organizations over the years. An outstanding example of this is the National Conference of Weights and Measures which has been in existence for over fifty years. In addition to running an annual conference, it also has standing committees which work on various problems.

The major result of the business meeting was the voting and formal acceptance of a resolution prepared and presented by Mr. Harvey Lance. This resolution specified that a standards laboratory conference should be held in Boulder, Colorado, in August 1962, thus breaking the ties with the Boulder Conference Committee which had organized the June 1960 Conference on Standards and Electronic Measurements and which was planning a similar Conference on Precision Electromagnetic Measurements in August 1962 at NBS Boulder. Another item approved as part of the resolution was the dissolution of the Ad Hoc Committee, replacing it with a continuing conference type of organization to be known as the General Committee for the National Conference of Standards Laboratories. This General Committee was charged with continuing and expanding the work begun by the subcommittees of the Ad Hoc Committee, arranging a continuing series of standards laboratory conferences, studying means of establishing a more formal organization of standards laboratories and initiating other activities as required to meet developing needs.

Formation of the General Committee of the NCSL

At the September 15, 1961 Standards Laboratory Conference, the Executive Committee of the Ad Hoc Committee appointed the following officers for the new General Committee of the NCSL.

Lloyd B. Wilson - Sperry Gyroscope Company (Chairman)
Charles E. Johnson - The Boeing Company (Vice Chairman)

Subsequently, in accordance with the resolution approved at the September 15 Conference, the Executive Committee of the Ad Hoc Committee elected the following additional officers and members of the General Committee:

Harvey W. Lance - NBS Boulder (Corresponding Sec.)
Charles E. White - Avco R.A.D. (Recording Sec./Treas.)
William G. Amey - Leeds & Northrup Co.
Herbert D. Barnhart - General Electric Co.
H. Curt Biggs - Sandia Corp.
Melvin Fruechtenicht - U. S. Army Ordnance
Jerry L. Hayes - U. S. Navy BuWeps
Wallace L. Horton - U. S. Air Force
Peter A. Joeschke - NAA Autonetics
William A. Wildhack - NBS Washington

The above members were chosen from representative phases of measurement standards work, such as users, suppliers, and standards laboratories, with fifty percent being definitely from standards laboratories.

Sponsors

Shortly after the September 1961 Conference, the National Bureau of Standards agreed to a request to be a sponsor of the NCSL. Solicitation of other sponsors is being held up pending clarification by the General Committee of the exact functions and responsibilities of sponsors of the NCSL.

Action Items vs. Organization Items--to Make NCSL a Going Organization

As I write this in early August, 1962, not quite a year has elapsed since the NCSL was established officially in Los Angeles on September 15, 1961. However, I think that we have had substantial progress during that time to make NCSL a going organization and to help standards laboratories solve some of their problems. During that time we have proceeded along parallel courses. One course was to get action items going as soon as possible to meet some of the needs of standards laboratories. The other course was to establish the organizational details over a period of a year or so, then to start work on action to meet the needs of standards laboratories. However, at the risk of being criticized for going off "half-cocked," action item work was started simultaneously with the development of the NCSL organization details.

Briefly, the organizational work of the past year consists of eleven committees set up to work on action items, plus two more concerned with organization and publicity. Also, a set of bylaws was prepared and revised several times. Several more committees should still be established, and further amendments to the bylaws will be required.

The need for action in a number of specific areas had been evolving gradually from the 1960 Conference on Standards and Electronic Measurements, the Ad Hoc Committee meetings, the Executive Committee meetings, and the September 1961 Standards Laboratory Conference. These needs seemed to fit into the list which is given below:

- (a) Measurement agreement and calibration traceability.
- (b) Evaluation, selection, and training of measurement standards personnel.
- (c) Standards laboratory work-load control.
- (d) Calibration procedures, techniques, and specifications.
- (e) Small standards laboratory operations.
- (f) Corporate standards laboratory organization and operation.
- (g) Design and construction of calibration rooms.
- (h) Calibration reliability and quality control.
- (i) Calibration cost reduction, cost justification, and value analysis.
- (j) Calibration terminology and definitions.
- (k) Terminology for use by instrument manufacturers in specifications and advertising claims.
- (l) Calibration equipment evaluation techniques to determine reliability of measurement standards and instruments.
- (m) Source material and recommended practices for organization and operation of standards laboratories.
- (n) Measurement standards information indexing, retrieval, and dissemination.
- (o) Improvement in the dissemination of information about NBS calibration services, techniques, and equipment.
- (p) Listing of calibration services available from standards laboratories other than NBS.
- (q) Standard environmental conditions for calibration; also other standard calibration conditions, such as voltage stability, impedance matching, etc.
- (r) Measurement standards and calibration equipment needs.
- (s) Means of anticipating new measurement demands on standards laboratories arising from new technology.

NCSL committees have been set up in many of the above areas to work towards solutions of the problems. The activities of six of these committees were pursued at six Standards Laboratory Management Workshops--three held in January 1962 at NBS Boulder, and three held in April 1962 at NBS Washington. Further work of these and other NCSL committees also took place in preparation for the August 8-10, 1962 Standards Laboratory Conference at the National Bureau of Standards in Boulder.

NCSL Committees and Committee Members

As stated previously, eleven committees were set up during this past year to work on action items, and two more were established for the areas of organization and publicity, respectively. The eleven action committees were set up as special committees. These special committees obtain basic source information by work of their committee members, from our Standards Laboratory Management Workshops, from our Standards Laboratory Conferences, from questionnaires, from literature and meetings of other people and groups, and by direct liaison with other organizations.

Special committees of the NCSL are expected to avoid duplicating the work of existing organizations in the fields of interest for NCSL. Instead, special committees are to cooperate with existing organizations by calling their attention to problems faced by standards laboratories and by requesting them to take whatever action they can to help solve these problems. In areas where the need is urgent and where suitable action cannot be obtained from other organizations or where the rate of effort by the other organizations does not meet our requirements, NCSL Special Committees should assist the other organizations or undertake solutions on their own.

Some of the types of efforts and results expected from the NCSL special committees are as follows:

- (1) Provide help in running Standards Laboratory Management Workshops, and assist in documenting results of the Standards Laboratory Management Workshops.
- (2) Work with program chairmen of NCSL Standards Laboratory Conferences to plan papers and/or sessions pertaining to the needs to be met by the respective committees.
- (3) Work with the NCSL Special Committee for Evaluation, Selection, and Training of Measurements Standards Personnel to provide training material in their respective subject areas.
- (4) Work with the NCSL Special Committee for Recommended Practices to prepare source material for standards laboratory practice manuals, as well as to recommend practices for standards laboratories.
- (5) Prepare a collection of sample forms and documents as used by various standards laboratories at the present time.
- (6) Prepare information on terminology and definitions currently in use by standards laboratories. Also, work with the NCSL Special Committee for Recommended Practices to prepare recommended terminology and definitions.
- (7) Work with the NCSL Special Committee for a Measurements Standards Information Center to prepare lists of references in the subject areas pertinent to the work of the respective committees. Also, work with the M.S.I.C. Committee to prepare general state-of-the-art information in the respective subject areas of each committee.

(8) Prepare a list of problems which still need to be solved in the subject areas of interest to the respective committees. Suggest solutions to be tried and evaluated.

Specific information about the various committees is given below, including a brief description of functions, the name of the chairman of each committee, and the latest information available to me regarding other members of the various committees.

Special Committee on Measurement Agreement and Calibration Traceability:

Chairman--S. C. Richardson, General Electric Co., Schenectady, N.Y.
Members --Herbert S. Ingraham, RCA, Camden, N.J.
Orville E. Kennedy, USAF, Newark, Ohio
Orval L. Linebrink, Batelle Memorial Institute, Columbus, Ohio
Kenneth G. Overbury, Sandia Corp., Albuquerque, New Mexico

Functions--To assist standards laboratories in making measurement agreement round-robins and audits. This includes round-robins in measurement categories where NBS does not provide required calibration services, as well as measurement audits to check and obtain the required degree of measurement compatibility in measurement categories where NBS does provide required calibration services. This also includes work to achieve accuracy ratios approaching 1:1 in the transfer of calibrations from NBS to primary type standards laboratories throughout the country.

Special Committee on Evaluation, Selection, and Training of Meas. Stds. Personnel:

Chairman--A. J. Woodington, General Dynamics/Astronautics, San Diego, Cal.

Functions--To take source material developed by NCSL committees, combine it with other suitable material as required, then put it into a more polished form suitable for use in the evaluation, selection and training of measurement standards personnel. The resulting material might be in the form of outlines, texts, slides, movies, programmed instruction texts or machines, etc. Suggest methods for both group training and training of individuals.

Special Committee on Standards Laboratory Work Load Control:

Chairman--Jerry L. Hayes, Navy BuWeps, Pomona, California
Members --Joseph M. Aldrich, Ryan, San Diego, Cal.
John R. Van de Houten, Aerojet-General, Sacramento, Cal.
Mike Rothbart, National Astro Laboratories, Pasadena, Cal.
Edwin P. Olejarczyk, General Electric, Utica, N.Y.

Functions--To study, organize, and disseminate information relating to improved methods for scheduling and controlling calibration facilities, personnel, and items to be calibrated so as to achieve short turnaround times consistent with requirements for low cost and for the maintenance of the required quality of calibrations performed.

Special Committee on Recommended Practices for Standards Laboratories:

Chairman--Kenneth G. Overbury, Sandia Corp., Albuquerque, New Mexico
Members --E. A. Anderson, Navy BuWeps, Pomona, Cal.
Mary Hoskins, Sheffield Corp., Dayton, Ohio
Orville E. Kennedy, USAF, Newark, Ohio
M. J. Leight, Hughes Aircraft Co., Culver City, Cal.
Frank D. Weaver, NBS, Boulder, Colo.

Functions--To provide basic source material, outlines, lists of references, and lists of terms and their definitions which might be useful to standards laboratories to help them establish or revise their operating practices. Also, to develop a handbook of recommended practices for standards laboratories.

Special Committee on Calibration Procedures, Techniques, and Specifications:

Chairman--Peter Joeschke, Autonetics, Downey, California

Functions--To devise means of supplying basic source information which standards laboratories can use in establishing or revising their own calibration procedures. Such source information might come from a calibration procedure information center, a calibration procedure exchange program, or from manufacturers of measurement standards and instruments. The information should provide material to assist standards laboratories in selecting proper calibration techniques, in writing the training portions of calibration procedures, and in specifying calibration points and conditions under which calibrations should be made to obtain desired accuracy levels. Help should also be provided in regard to calibration procedure formats and suggestions for simplifying the writing of calibration procedures.

Special Committee on Calibration Cost Reduction and Value Analysis:

Chairman--Herbert D. Barnhart, General Electric Co., Syracuse, N.Y.

Functions--To apply some of the currently popular and proven methods of cost reduction and value analysis to standards laboratory operations and to project this to the overall effect on company products and the national standards program. Also, to study and suggest methods of accounting which will permit recognition by management of hidden costs of poor calibrations--e.g., things such as rejects or rework which often can be traced to inadequate measurements and calibrations. In addition, to study means of using cost reduction and value analysis techniques to determine *realistically* the need for procurement of standards laboratory equipment, personnel, calibration rooms, etc.

Special Committee on Corporate Standards Laboratory Organization and Operation:

Chairman--Lewis Wallace, IBM, Kingston, N.Y.

Functions--To study techniques by which measurement standards and calibration laboratories can assist the National Bureau of Standards in providing primary-level calibration services and improving the overall compatibility of measurements and standards throughout the United States. Also, to devise means by which large companies can reduce costs in providing their own calibration services by making use of either a central standards laboratory or specialized individual standards laboratories in various parts of the company. In addition, to study means by which corporate standards laboratory techniques can be applied to smaller companies and to heterogeneous groups of standards laboratories.

Special Committee on Reliability of Measurement Standards and Instruments:

Chairman--Leon Hachey, Hughes Aircraft Co., Fullerton, Cal.

Members--Leon Dean, Dean Laboratories, Mansfield, Mass.
R. F. Estoppey, Weston Instrument Co., Newark, N.J.
Robert P. Heckelmann, Sperry Gyroscope Co., Great Neck, N.Y.
Harry Pegg, Leeds & Northrup Co., Philadelphia, Pa.
Harry S. Pyle, McDonnell Aircraft Co., St. Louis, Mo.
Len Silbert, Thiokol Chemical Corp., Huntsville, Alabama

Functions--To develop techniques which can be used by standards laboratories for evaluation of the characteristics of measurement standards and instruments to determine special characteristics or to determine conformance with manufacturers' specifications. Also, to study vendors' specifications for specific types of measurement standards and instruments in an effort to develop uniform specifications and terminology which can be used by vendors to provide better descriptions of the measurement standards and instruments which they sell.

Special Committee on NBS Calibration Service Information:

Chairman--Howard S. Johnson, Martin Marietta Corp., Denver, Colo.

Members--W. F. Snyder, NBS, Boulder, Colo.
W. Reeves Tilley, NBS, Washington, D.C.
Evan Lapham, Avco R.A.D., Wilmington, Mass.

Functions--To study requirements of standards laboratories for improved dissemination of information regarding the calibration services provided to them by the National Bureau of Standards. This includes the type of information, format in which the information is presented, provisions for keeping it up to date by necessary additions, deletions, and changes in services, etc. Recommendations should be made to NBS regarding changes in its calibration service information as indicated by the studies described above.

SPECIAL NOTE: THIS COMMITTEE COVERS ONLY NBS CALIBRATION SERVICE INFORMATION, NOT CALIBRATION TECHNIQUES OR CALIBRATION EQUIPMENT UNDER DEVELOPMENT OR IN USE AT NBS. I THINK THAT, ALTHOUGH THERE IS A RELATIONSHIP HERE WITH THESE OTHER NBS ACTIVITIES, THERE IS ENOUGH TO BE DONE IN THIS NBS CALIBRATION SERVICE INFORMATION AREA ALONE TO JUSTIFY HAVING A SEPARATE COMMITTEE FOR NBS TECHNIQUES AND EQUIPMENT DEVELOPMENT INFORMATION, AS DESCRIBED BELOW.

Special Committee on NBS Techniques and Equipment Development Liaison:

Chairman--Evan Lapham, Avco R.A.D., Wilmington, Mass.

Functions--To study requirements of standards laboratories for improved dissemination of information regarding calibration techniques and equipment under development or in use at the National Bureau of Standards. Also, to work with NBS to provide necessary information as disclosed by the above studies. The principal requirements here are to provide speedier access to such information, to provide more of it, and to make it easier to acquire. We recognize that NBS already has certain formal channels for providing such information, and in addition NBS personnel have always been extremely cooperative in providing information to people making visits or writing letters. However, these methods supply only a sampling of the information really needed. The best way of obtaining such information appears to be through personal visits to NBS, but these are awkward, time consuming, and expensive for the people who seek the information; in addition, this diverts NBS personnel from important research activities which they should be doing to provide improved national standards and calibration services which are so badly needed today in many measurement categories.

Special Committee for Measurement Standards Information Center:

Chairman--Lloyd B. Wilson, Sperry Gyroscope Co., Great Neck, N.Y.

Members--Thomas R. Hamilton, Lockheed Missile and Space Co., Sunnyvale, Cal.
William L. Vandal, Aeroneutronic, Newport Beach, Cal.
Charles E. Stone, RCA Service Co., Patrick AFB, Florida

Functions--To work with other organizations to study the need for and perhaps to help establish a measurement standards information center. Also, to develop techniques and establish services for the storage, retrieval, and dissemination of information from NCSL committees and from other sources relating to the organization, operation, and evaluation of measurement standards and calibration laboratories. In addition, to work with other NCSL committees to help them acquire and organize information in such a way that it can be more easily indexed, retrieved, and disseminated.

There are also two standing committees to help in planning and operating the NCSL. These committees are as follows:

Standing Committee on Organization:

Chairman--William Wildhack, NBS, Washington, D.C.

Members--Charles White, Avco R.A.D., Wilmington, Mass.
Lloyd B. Wilson, Sperry Gyroscope Co., Great Neck, N.Y.

Functions--(1) To prepare a basic set of bylaws for approval by the General Committee, and to suggest amendments from time to time, as required.
(2) To work with the various NCSL committees to prepare activity guides to cover their organization and operation.
(3) To study means of improving the organization and operation of NCSL to meet better the needs of measurement standards and calibration laboratories.
(4) To study means of establishing a more formal organization of standards laboratories, as called for in the resolution approved at the September 15, 1961 Standards Laboratory Conference in Los Angeles.

Standing Committee on Publicity:

Chairman--Charles White, Avco R.A.D., Wilmington, Mass.

Members--John Orme, Aerojet-General Corp., Azusa, Cal.

Functions--(1) To provide general types of publicity information regarding NCSL and its activities.
(2) To supply necessary publicity information to magazines, newspapers, etc.; also, to coordinate the supplying of such information by others within NCSL.
(3) To provide NCSL's own publicity media, such as periodic newsletters or publicity releases.
(4) To provide or coordinate the providing of formal answers, where required, for questions or problems concerning the public relations of NCSL.

The Standards Laboratory Management Workshops

One of our major action items during the past year, in addition to the initiation of work by the various NCSL special committees, was the holding of six Standards Laboratory Management Workshops. These workshops were really extensions of the activities of some of the NCSL special committees; each of the six workshops corresponded to the name and functions of one of the NCSL special committees. In this way, the workshops provided a start for those committees by initiating discussions of problems, current practices, and suggestions for further work of the respective committees. In addition, each committee chairman had the opportunity to solicit interested and competent people to serve with him on his committee.

Of course, the activities and discussions at the workshops helped in other ways too. Individuals who attended the workshops gained information which helped them compare their standards laboratory operations with those of other organizations. The attendees were also able to gain information which should help them to deal more effectively with problems of standards laboratory operations in their own organizations.

Attendance at these workshops was by invitation only, in an attempt to keep the size of each workshop down to a reasonable working level. Even then, attendance was more than we had anticipated; attendance at each workshop was in the range of 50 to 60 people. In issuing the invitations we tried to get a representative sampling of people from a variety of companies and organizations, and also from a number of different parts of the United States. We succeeded quite well in both of these objectives, with attendees coming from private industry, NBS, Army, Navy, Air Force, and the Atomic Energy Commission.

The list of Standards Laboratory Management Workshops held to date is as follows:

Workshop No. 1--CALIBRATION TRACEABILITY AND MEASUREMENT AGREEMENT
Tuesday, January 23, 1962; Chairman--S.C. Richardson, General Electric Co.

Workshop No. 2--EVALUATION, SELECTION, AND TRAINING OF MEAS. STDS. PERSONNEL
Wednesday, January 24, 1962; Chairman--A. J. Woodington, GD/Astronautics

Workshop No. 3--STANDARDS LABORATORY WORK LOAD CONTROL
Thursday, January 25, 1962; Chairman--Jerry L. Hayes, Navy BuWeps

Workshop No. 4--CALIBRATION PROCEDURES, TECHNIQUES, AND SPECIFICATIONS
Monday, April 16, 1962; Chairman--Peter Joeschke, Autonetics

Workshop No. 5--RELIABILITY OF MEASUREMENT STANDARDS AND INSTRUMENTS
Tuesday, April 17, 1962; Chairman--Leon Hachey, Hughes Aircraft Co.

Workshop No. 6--CALIBRATION COST REDUCTION AND VALUE ANALYSIS
Wednesday, April 18, 1962; Chairman--Herbert Barnhart, General Electric Co.

Workshops Nos. 1, 2, and 3 were held at the Boulder Laboratories of the National Bureau of Standards. Workshops Nos. 4, 5, and 6 were held at NBS in Washington, D.C.

In my opinion, these workshops were quite valuable, and I think that more should be held in the future. There are still several NCSL committees which could benefit by having workshops run in their areas of interest, and in some cases workshops might also be run in special areas of interest which are not covered by any present NCSL committee or which are just part of the area of interest of a particular NCSL committee.

The 1962 Standards Laboratory Conference

This conference, to be held August 8-10, 1962 at the Boulder Laboratories of the National Bureau of Standards, is the focal point of our action item work for the past year. The general purpose of this Conference is to provide a medium for disseminating information on the organization and operation of measurement standards and calibration laboratories with the goal of promoting increased competence, better organization, and uniform practices among the Nation's standards laboratories. More specifically, reports will be given by NCSL committees, with various committee chairmen, committee members, and others presenting papers or participating in panel discussions.

At the time this is written, in early August of 1962, it appears that this conference should be very successful. One indication of this is the excellent program planning done by Charles Johnson and the session chairmen. Another indication is that, at this time, the registration is well above the 500 mark.

Personal Observations and Recommendations

The following five things stand out in my mind at this time regarding past and possible future operations of the NCSL.

(1) WE MUST ASSURE CONTINUITY IN THE ENTIRE OPERATIONS OF THE NCSL. Many of the things I have described in this report are just starting to produce results, and some will require continuing effort for a long time. Interruptions and discontinuities are inevitable in any organization, but we have not yet achieved the state of maturity where we can afford the luxury of having too many such changes. One way to achieve continuity, and one which I strongly recommend, is to establish a full-time secretariat for NCSL. Another way is to try to work out a logical progression system for people doing committee work and for members of the General Committee.

(2) WE MUST SPREAD OUR BASE OF ACTION. This is necessary to bring in new blood, to help us become more democratic, and to get extra help on the many problems which standards laboratories face. I have already detected signs of "saturation" in the efforts of some individuals, including myself. This does not imply a lack of interest or a lack of willingness to do extra work for the NCSL. It just means that each of us has a limit to how much we can do in our job, our home life, and in outside activities such as NCSL. However, we definitely need more help to get things done that have to be done, and to get them done within a reasonable period of time. Paradoxically, there probably are many people who would like to help out more, but our problem here is getting to know them and then matching them up to NCSL work which would be of interest to them and where they would be competent to do it. We can profitably explore this approach further. Another approach which we are just starting to develop is the matter of liaison and cooperation with other organizations.

(3) WE MUST BE FLEXIBLE AND FAST ACTING. There is an old saying that "time solves all problems." However, time may turn out to be our greatest enemy if we wait for it to solve our standards laboratory problems for us. To be flexible and fast acting will require a combination of the best features of both the small organization and the large organization. What this really means is that we must be able to function like a small organization in terms of independence and speed of action to meet the needs of measurement standards and calibration laboratories. However, at the same time we must also be able to bring to bear the specialized services and talents which only large professional societies and trade organizations now seem to have for their operations. This seems to create a paradoxical situation, but I think that it can be resolved, once we

recognize it and admit that it is a problem.

(4) WE SHOULD BE PROBLEM ORIENTED. Many organizations--trade associations, professional societies, and the like--appear to be little more than social clubs or pseudo-debating societies. They run meetings and publish journals more for the sake of meeting a fixed time schedule than to satisfy any real needs. When deadlines approach, they frantically try to find someone to speak or write an article "on some subject of interest to our members." I will not deny that this does occasionally produce some useful results, but it appears to me to be very inefficient. What we need in NCSL is to focus our meeting and publication efforts on specific problems faced by standards laboratories; if there are no problems in a particular area, we should forego having the meeting or attempting to publish anything on the subject. The fact that we are "laboratory" oriented should help us to more readily determine the problems to be solved. At the same time, our efforts to solve such problems will help not only the standards laboratories but also the individuals who work in those standards laboratories.

(5) WE MUST HELP STANDARDS LABORATORIES TO HELP THEMSELVES. As a national organization we can do much for standards laboratories through the cooperative efforts of our members--actually more than any one or a few laboratories could hope to do on their own. For example, this includes things like determination of uniform practices, terminology and definitions, conducting measurement agreement operations, or running a calibration procedure exchange program. However, we should also do everything we can to help standards laboratories do things themselves "in-house." Here the examples include such things as helping them write better calibration procedures by suggesting good formats and supplying source material, or helping them to do their own training by supplying course material and telling them how to teach their own formal courses or how to use on-the-job training. Another thing they should be helped to do is to police their own standards laboratory operations with the help of source material and recommended practices prepared by our Recommended Practices Committee.

Our efforts and successes to date owe much to the splendid cooperation and assistance received from the National Bureau of Standards--particularly Dr. Astin, William Wildhack, Harvey Lance, and Reeves Tilley. I also want to thank Charles Johnson who is the program chairman for our 1962 Conference, Mr. James Brockman of NBS Boulder, all of our committee chairmen, and all of those people who have supported our activities by work on committees and attendance at our workshops and our forthcoming 1962 Standards Laboratory Conference.

Lloyd B. Wilson,
August 3, 1962 Chairman, General Committee

Session 3. NCSL Business and Information

PANEL ON STANDARDS LABORATORY INFORMATION DISSEMINATION

Moderator:
Lloyd B. Wilson, Sperry Gyroscope Co.
Great Neck, New York

Members:
Herbert D. Barnhart, General Electric Co.
Syracuse, New York

Leon E. Hachey, Hughes Aircraft Co.
Fullerton, California

Thomas R. Hamilton, Jr., Lockheed Missiles and Space Co.
Sunnyvale, California

Peter A. Joeschke, Autonetics
Downey, California

Howard Johnson, Martin-Marietta Co.
Denver, Colorado

Evan G. Lapham, AVCO R. A. D.
Wilmington, Mass.

Charles E. Stone, RCA Service Co.
Patrick AFB, Florida

Summary

Mr. Wilson: This session will contain reports of the NCSL committees not represented in other sessions. It also will emphasize the information dissemination aspects of the work of these committees and of NCSL in general.

Today we are faced with an information explosion which is creating serious retrieval problems and which will cause worse problems in the future if realistic steps are not taken to cope with it. NCSL has an active interest in this problem because we not only acquire information (through this Standards Laboratory Conference, through the Standards Laboratory Management Workshops, and through various NCSL committees), but we also disseminate it.

The various phases of the information handling problem include acquisition, processing, indexing, storage, retrieval, and dissemination. These phases are illustrated in figure 1, which depicts the operation of an information center. This chart is from a paper titled, "The Management of Scientific Intelligence" by G. S. Simpson, Jr., of Battelle Memorial Institute and is reproduced by permission. An information center has been described as "a new organization designed not to replace, but to supplement the library" and also as "a link between the flood of technical information located in the library, and the scientist or engineer."[1] NCSL should go one step beyond the traditional professional society approach and pay particular attention to the acquisition phase. It has been suggested[1] that in addition to preparing abstracts, authors should classify their papers by indicating appropriate categories or descriptors. The author is most familiar with the content of his paper and should learn to classify it in some standard way. Quite a bit of work has already been done by others in this field. One book gives a comprehensive listing of science information services in the United States.[2] This book includes brief descriptions of the information service activities of 427 different organizations or projects, with an average of slightly more than one page devoted to each. There are still other information centers. For example, there is the Electronic Component Reliability Center at the Battelle Memorial Institute. The interests of this center may parallel the interests of our NCSL Special Committee on Reliability of Measurements Standards and Instruments.

In NCSL most of our information center work is in the planning stage. Our Measurements Standards

---

[1] S. Isaacson, "Industry Faces the Information Challenge", presented October 10, 1961, at a session titled "Technical Information; The Paper Curtain" at the Annual Meeting of the American Rocket Society, New York City.

[2] Specialized Science Information Services in the United States, National Science Foundation Publication No. NSF 61-68, November, 1961. Published by the Superintendent of Documents, U.S. Government Printing Office, Washington 25, D.C.

Information Center Committee is the newest of the NCSL committees and was organized only in April of this year. As already mentioned, NCSL is acquiring information in several ways, but its work at present does not involve the more sophisticated types of information indexing and retrieval. Several NCSL committees have as their chief function the organizing and processing of information on such subjects as recommended practices for standards laboratories, and evaluation, selection, and training of measurements standards personnel. It is expected that the Committee also will become active in organizing and processing information.

NCSL Special Committee on Calibration Cost Reduction and Value Analysis

Herbert D. Barnhart, Chairman

Objectives: To apply some of the currently popular and proven methods of cost reduction and value analysis to standards laboratory operations and to project this to the overall effect on company products and the national standards program. Also, to study and suggest methods of accounting which will permit recognition by management of hidden costs of poor calibrations--e.g., things such as rejects or rework which often can be traced to inadequate measurements and calibrations. In addition, to study means of using cost reduction and value analysis techniques to determine *realistically* the need for procurement of standards laboratory equipment, personnel, calibration rooms, etc.

Mr. Barnhart: With the increased measurement accuracy required of companies doing government contract work, the cost of this type of work has become of more concern to management. The Committee on Calibration Cost Reduction and Value Analysis was formed to aid management in applying cost reduction and value analysis techniques to achieve more efficient instrumentation and measurement programs. The committee sponsored a one-day Standards Laboratory Management Workshop at the National Bureau of Standards in Washington, D.C., on April 18, 1962. Information presented at the workshop indicated that many companies have cost reduction and value analysis programs directed toward manufactured hardware, but that very little effort from either of these programs was directed toward instrumentation operations. There were quite a few indications that measuring instruments are being purchased which have excess performance capabilities and that unnecessary costs are being incurred in providing calibration services for such instruments. Cases were noted in which design accuracies probably in excess of needs were specified for certain parts of overall systems, while at the same time inadequate accuracy was specified for other parts of the overall system. The committee wants to devise a technique for looking at the overall system so that the correct accuracy can be specified and achieved in each part of the system and throughout the calibration program. The approach of the committee will be to attempt to find means of applying the concepts used in hardware value analysis programs to calibration and measurement value analysis programs. This may be difficult. At a forthcoming workshop it may be possible to facilitate the development of techniques by working on a specific example. The committee has initiated the following program:

(1) Organizing information to aid in assigning correct costs to each calibration program, to aid in determining whether equipment is being over-calibrated or under-calibrated.
(2) Formulating plans to aid in analyzing each instrument requirement needed to perform a system calibration, so that the accuracy of each individual instrument can be made compatible with all other instruments in the system and so that the required system accuracy can be achieved.
(3) Studying the intangibles involved in performing instrument calibrations for a manufacturing effort, including the effect of calibrations on the level of engineering effort required and on the compatibility of equipment on the assembly line.
(4) Obtaining detailed cost figures on all elements of calibration systems to enable companies to spot costs which are out of line.

Special Committee on Reliability of Measurement Standards and Instruments

Leon E. Hachey, Chairman

Objectives: To develop techniques which can be used by standards laboratories to evaluate the characteristics of measurement standards and instruments, to determine special characteristics, or to determine conformance with manufacturers' specifications. Also, to study vendors' specifications for specific types of measurement standards and instruments in an effort to develop uniform specifications and terminology which can be used by vendors to provide better descriptions of the measurement standards and instruments which they sell.

Mr. Hachey: The committee has decided initially to direct its activities toward improving

(1) uniformity and realism of instrument specifications and
(2) uniformity of instrument evaluation methods.

The committee sponsored a Standards Laboratory Management Workshop at the National Bureau of Standards, Washington D.C., on April 17, 1962. As an example of the need for the work which the committee is undertaking, manufacturers of precision potentiometers list accuracy in at least 16 different ways. Work has begun on recommended specification forms for the following six types of instruments: potentiometers, d-c bridges, precision decade resistors, d-c voltage dividers, a-c voltage dividers, and LRC bridges. First drafts on three of these are already in circulation for comments and suggestions.

The committee does not intend to confine its work to electrical instruments but rather to cover all categories of instruments. The listed instruments were chosen because they are in common use and because the manufacturers themselves have provided a start on uniformity of specifications. Work on additional categories of instruments will be begun as interested volunteers for committee work are obtained.

The committee is using the following procedures:

(1) A committee member searches vendor literature for information on a particular type of instrument. He need not include every vendor but, should include enough to show norms and deviations.
(2) A chart is prepared which compares the specification forms of the major vendors.
(3) A recommended form is drawn up, using the best features from the chart and from other sources.
(4) A draft copy of the chart and of the recommended form is sent to each member of the committee for comments and suggestions.
(5) A second draft is drawn up utilizing the suggestions of the committee members whenever applicable. A copy is sent to the vendors involved for their comments and suggestions. The form is also distributed to those committee members who wish to participate. A third draft of the form is then made.
(6) A portfolio containing the background information and the recommended (third draft) is sent to the committee chairman.
(7) This form is reviewed by the committee chairman and sent to the chairman of NCSL for adoption as a recommended form. The method of adoption is yet to be decided.

No one will be required to use this form; however, by the time it has reached this stage it will have a backing of a large number of measurements people throughout the United States. This fact alone should do much to promote uniform and reliable vendors specifications. The committee solicits information from all who are already doing instrument evaluation work regarding the types of forms that are used, the procedures, the method of selecting, instruments to be evaluated, the method of evaluation, and the type of report issued.

Special Committee on Calibration Procedures, Techniques, and Specifications

Peter A. Joeschke, Chairman

Objectives: To devise means of supplying basic source information which standards laboratories can use in establishing or revising their own calibration procedures. Such source information might come from a calibration procedure information center, a calibration procedure exchange program, or from manufacturers of measurement standards and instruments. The information should provide material to assist standards laboratories in selecting proper calibration techniques, in writing the training portions of calibration procedures, and in specifying calibration points and conditions under which calibrations should be made to obtain desired accuracy levels. Help should also be provided in regard to calibration procedure formats and suggestions for simplifying the writing of calibration procedures.

Mr. Joeschke: This committee sponsored a Standards Laboratory Management Workshop at the National Bureau of Standards, Washington, D.C., on April 16, 1962. During this workshop it became apparent that several fundamental questions need to be answered in order to guide the committee in its work. For example, what is a standards laboratory in terms of NCSL? Further, does a standards laboratory just calibrate or does it also so repair work? The answers to these and similar questions will influence the direction of efforts of the committee. The workshop also made it clear that many problems related to calibration procedures and techniques also are closely related to individual company policies and must be solved by the individual companies. Early in the year, questionnaires were sent to 50 major calibration laboratories to get a cross-section of views on calibration procedures. Thirty returns were received, revealing the following information and opinions:

(1) "Calibration Procedures" is the most common name for documents which tell someone what and how to calibrate.
(2) The prime source of usable calibration procedures is the "in house" writing effort.
(3) Next in importance as sources are the manufacturers of instruments.
(4) The most significant reason for having calibration procedures is to aid in the standardization of techniques and calibration equipment.
(5) The content of calibration procedures is governed mainly by product requirements and by the capability of the individual laboratory.
(6) Procedures for all calibrations should be on file.
(7) It is not advisable to substitute skill for written procedures except on some production-type setups.
(8) The cost of in-house written calibration procedures ranges from $40 to $5,000, with an average cost of $489.

The last point raised on the questionnaire was whether companies and government organizations would participate in a calibration procedures exchange program. Ninety-five percent of the replies indicated a willingness to do so. Several groups already are engaged in calibration procedure exchanges. About a year ago the metrology laboratory of one large company initiated such a program among the divisions of the company, and this has resulted in savings to the company in excess of $50,000. The success of this program suggests that the concept could be carried out on a national basis, somewhat as follows: On the basis of mutual agreement among participants, each participant in the exchange program issues a list of calibration procedures to the fellow participants. Each participant selects from those lists the procedures he would like to obtain. The participant originating the procedures then provides one copy of each procedure requested.

The committee also is concerned with improving the calibration procedures provided by instrument manufacturers. It intends to request associations of instrument manufacturers to cooperate with NCSL in regard to the format and content of their procedures.

Special Committee on NBS Calibration Service Information

Howard Johnson, Chairman

Objectives: To study requirements of standards laboratories for improved dissemination of information regarding the calibration services provided to them by the National Bureau of Standards. This includes the type of information, the format in which the information is presented, and provisions for keeping it up to date by necessary additions, deletions, and changes in services, etc. Recommendations should be made to NBS regarding changes in its calibration service information as indicated by the studies described above.

Mr. Johnson: This committee was formed as the result of needs expressed at the third Standards Laboratory Management Workshop, which was held at the NBS Boulder Laboratories on January 25, 1962. There the feeling was expressed that factual information on NBS services was not available in a single up-to-date and easily understandable document that could be used readily by standards laboratories. It also was felt at the workshop that the accuracy statements provided by NBS did not fully meet the needs of standards laboratories. One result of the work of the committee was a compilation of sources of information on NBS calibration services. Preprints of this compilation were distributed at the Conference, and the compilation was published later in the NBS Technical News Bulletin.[3] Partly at the suggestion of the committee, the editors of the Technical News Bulletin plan to include additional items of interest to standards laboratories.

Special Committee on NBS Techniques and Equipment Development Liaison Chairman

Evan G. Lapham, Chairman

Objectives: To study the requirements of standards laboratories for improved dissemination of information regarding calibration techniques and equipment under development or already in use at the National Bureau of Standards. To provide NBS with information regarding these requirements. The principal requirements are believed to be more information and speedier and easier access to the information.

Mr. Lapham: The field of interest of this committee is technical liaison with NBS but excludes matters related to calibration service information, which are covered by Mr. Johnson's committee. The committee is investigating means of facilitating the transfer of information on NBS techniques and equipment to those who want and need it. To date, work of the committee has been directed toward determining presently available sources of information. A compilation of information on this subject was passed out during the conference. (A similar, though earlier, compilation is available from the Government Printing Office.[4]) Attention was called to the NBS Journal of Research, the NBS Technical News Bulletin, and particularly to the three volume NBS Handbook 77 on Precision Measurement and Calibration.[5] This handbook brings together a whole series of papers published over many years on the subject of precision measurement. Attention also was called to the report at an earlier session that NBS has tentative plans to give seminars of perhaps three days to two weeks in duration, covering selected specialized fields. It was noted that the Electricity Division of NBS Washington has been working up papers that contain information in some detail on calibration procedures and are not just research papers on principles. One of these on direct current resistance apparatus already has appeared[6] and others are in preparation.

Special Committee on Measurements Standards Information Center

Lloyd B. Wilson, Chairman

Objectives: To work with other organizations to study the need for and perhaps to help establish a measurement standards information center. Also, to develop techniques and establish services for the storage, retrieval, and dissemination of information from NCSL committees and from other sources relating to the organization, operation, and evaluation of measurement standards and calibration laboratories. In addition, to work with other NCSL committees to help them acquire and organize information in such a way that it can be more easily indexed, retrieved, and disseminated.

Mr. Wilson: In fulfillment of its objective of working with other organizations, the committee has discussed its interests in a preliminary way with the ISA. It has also requested information from the Engineers Joint Council, which has an action plan and several committees active in this area. In addition, the committee has been making a survey to determine how other information centers are organized and operated. Two studies are under way by subcommittees.

Report of Subcommittee No. 1, by Thomas R. Hamilton, Jr.: The subcommittee has found that in some cases the problem is not the lack of information, but rather that there is too much information and consequently there is great difficulty in sorting it out. This is one factor with which a measurements standards information center would have to contend. In order to save time and reduce redundancy, such a center might maintain a cross index of calibration procedures, an index of standards laboratories, a bibliography of technical articles, a listing of contractural requirements, and possibly a service for answering individual questions. The file and cross index of procedures could include procedures of the military services, the National Bureau of Standards, and industrial laboratories. Access to such a file could prevent duplication of effort in the preparation of similar procedures by other companies. The index of standards laboratories could list the areas of measurement in which the laboratory was working, the reference standards used by the laboratory, the availability and costs of calibration services, and the qualification of the laboratory to perform

---

[3] Standards and Calibration, Sources of Information on NBS Services, NBS Technical News Bulletin, Vol. 46, No. 9, (Sept. 1962).

[4] U.S. Department of Commerce--Part III, National Bureau of Standards, October, 1960, No. 8, (NSF 60-59), Scientific Information Activities of Federal Agencies, National Science Foundation, Washington 25, D.C.

[5] Precision Measurement and Calibration, Vol. I--Electricity and Electronics; Vol. II--Heat and Mechanics; Vol. III--Optics, Metrology, and Radiation; NBS Handb 77 (Feb. 1, 1961).

[6] Paul P. B. Brooks, Calibration Procedures for Direct-Current Resistance Apparatus, NBS Mono. 39, March 1, 1962, Superintendent of Documents, U.S. Government Printing Office, Washington 25, D.C.

such services. Such knowledge could minimize transportation, time delays, and cost, and could reduce the calibration uncertainties caused by the long distance transporation of standards. An indexed bibliography of papers, books, and articles on calibration would be extremely useful. There is an ever-increasing quantity of general and specific military specifications. A centralized listing of such requirements would permit more effective dissemination of information on these. The answering service for individual questions would include assistance with measurement problems, with new instruments on the market, etc. It will cost money to implement these five recommendations, but the cost may well be less than what already is being paid in time and money for this information. NCSL should determine the kinds of information most urgently required by its members and should provide this information to the membership. A questionnaire has been prepared which should provide this type of information to the committee chairman.

Report of Subcommittee No. 2, by Charles E. Stone: Although conferences are an excellent means of communication in the measurements standards field, it is important to continue communication on a day-to-day basis and to find methods of eliminating barriers to effective communication. These barriers consist of such things as distance, time, proprietary information, lack of knowledge of sources of information, and inter-service security restrictions. However, these all can be overcome. A recent study has indicated, for example, that the exchange of information among the three national missile ranges can be improved with a minimum effort on the part of each range standardization group and with a great potential gain. This study suggests an approach that can be followed in establishing a measurements standards information exchange and dissemination system in industry. A central office should maintain files listing on standard format such information as (1) the field of measurements in which work is being carried on and the level of accuracy at which the work is being done, (2) the name of the organization or company, (3) the names of the management people in charge of the work, (4) an indication of the level of proprietary restrictions imposed by the company, (5) a delegation of authority to the clearing house to circulate information contributed by the company, and (6) an agreement to contribute copies of direct mail communications on the subject to the central file. Then any interested organization could find information on other organizations doing work of interest and could establish proper direct communications. In addition, it could get pertinent information from the central file. The organization originating the information would retain control over its release. Management of such an operation seems appropriate for NCSL.

Mr. Wilson: A measurements standards information center could provide basic state-of-the-art information, process the outputs of NCSL committees, and operate a clearing house for the exchange of information. In the first category are included bibliography lists, abstracts, standard forms used in operating a standards laboratory, and full-length documents. The second category includes principally how-to-do-it information on training, calibration procedures, standards laboratory work control, the quality and reliability of measurements standards, recommended practices, and other NCSL committee outputs. In the third category, the clearing house might operate calibration procedure exchanges, provide measurements standards information by mail, disseminate results of evaluations on the quality and reliability of measurement standards and instruments, and produce and distribute training manuals. Thus, such a center would need to have adequate facilities for printing and mailing as well as competent technical guidance.

Discussion Period

Question. "We had a problem in microwave measurements and asked our research library to make a literature search for us. We received a list of 700 references on this particular subject and were completely overwhelmed. Is this typical of the service we might expect from an information center?"

A good information center would have a suitable fine structure in its indexing and retrieval system to avoid getting too many references of a general nature. It also would provide abstracts and where necessary would provide complete copies of documents. In other words, instead of ending up with a list of 700 references you might get only the 5 or 10 that were particularly pertinent to your problem.

Question. "As organizations grow older they tend to get very highly organized and develop a sort of 'hardening of the arteries' in their information dissemination activities. What precautions will NCSL apply in order to avoid similar difficulties?"

At this stage no formal precautions have been worked out. At least an effort will be made to avoid these difficulties, but only time will tell how successful it will be.

Question. "With free exchange of information through an information center, how does one arrive at general agreement in any particular area?"

Arriving at a general agreement in any particular area of NCSL is largely a problem of committee work. Once the results of committee work are reviewed and approved, they will be disseminated through the measurement standards information center and, in some cases certainly, through the output of the Recommended Practices Committee. Before the general agreement stage has been reached, NCSL can fulfill a useful function by the dissemination of basic source material. This is information concerning the present state of the art as it is being practiced, and it is up to the individuals receiving this source material to use it as they see fit.

Question. "What is a standards laboratory in terms of NCSL?"

The answer to this question can be found in the NCSL bylaws. The term "measurement standards and calibration laboratories" includes all such laboratories performing any category of calibration work: dimensional, physical, electrical, or chemical, and at any level of calibration, including the calibration of instruments and product items.

Question: "Does a standards laboratory just calibrate or does it repair as well?"

The definition of a standards and calibration laboratory quoted above was arrived at after considerable discussion, but the repair aspect was not considered in this definition. This is a suitable subject for later consideration by the Recommended Practices Committee. As far as actual present day practices are concerned, some laboratories do no calibration and repair. The standards laboratories, the lower level calibration laboratories, and the repair and maintenance activities often are separate but closely coordinated activities. In other companies, anything that normally is calibrated in the primary standards laboratory will also be repaired if necessary by that laboratory. In the lower echelons it is impossible to separate the two types of work.

Question. "If you plan to describe NBS techniques and equipment, you may find that some of the equipment in NBS calibration setups are not as suitable for use in NCSL laboratories as is alternate equipment in use elsewhere. How will you learn about this alternate equipment and how will you arrive at a recommendation to the NCSL members?"

The purpose of the Special Committee on NBS Techniques and Equipment Development Liaison is to, as far as possible, keep NBS informed of the needs for various types of information, with the hope that information may be made available sooner than otherwise would be the case. The accuracy and completeness of the information is an NBS responsibility and it will be given out in NBS publications as usual--not in NCSL publications. In other words, this committee is really concerned only with liaison with NBS and not with alternative sources of information or ways of doing things. In regard to the last part of the question, "how will you arrive at a recommendation to the NCSL members", the objectives of the committee are not to recommend but to provide information.

Question. "Are you attempting to write actual specifications on each specific type of instrument or only to establish definitions of terms which are necessary in establishing characteristics?"

Eventually both may be done. However the initial work of the Committee on Reliability of Measurement Standards and Instruments has been along the lines of working out suitable specifications for certain types of instruments. This is essentially an attempt to formulate a set of performance requirements that an instrument should meet if it is to belong to a certain accuracy class.

Question. "A lot has been said about the responsibility of NBS in providing information on research, calibration services, etc., but how about the responsibility industry has of informing NBS of its medium and long range requirements for measurement services?
What action does NCSL plan in this regard?"

NCSL has no specific plans in this area, but it is part of the general objective of the organization to determine such needs. For instance, our bylaws state: "The NCSL undertakes to identify measurement areas in which adequately reproducible standards or techniques of measurement are lacking and to encourage research and development activities directed toward improvement of such standards or techniques." A useful function of the proposed measurements standards information center would be to acquire information about needs for improved measurement and calibration services, to analyze and index the information, and to pass it on to NBS.

Question. "Does the scope of your activity include committee evaluation of actual instruments or just the establishment of evaluation techniques and standard forms?"

NCSL has no intention of doing any actual evaluation of instruments. It will attempt to suggest and recommend forms and methods and techniques for evaluation so that if instruments are being evaluated, different laboratories will do the evaluation in a similar manner. The information resulting from such evaluations then will be comparable and can be gathered and summarized properly by NCSL. This type of work must be handled carefully because of the possibility that an instrument manufacturer might take legal action if he thought NCSL was disseminating information derogatory to his particular product. The approach used in similar situations such as that of the Electronic Component Reliability Center at Battelle Memorial Institute should be investigated in this regard. That Center provides evaluation information on components. The AIEE Committee on High Frequency Measurements has a subcommittee which is actively working on the standardization of specifications and on recommended test procedures for verifying these specifications. NCSL may be interested in following the work of this committee.

Question. "This morning the subject of NBS services was covered in greater detail than is available in the literature. Are preprints of this morning's talk available?"

Preprints are not available, but it is intended to include these papers in the Proceedings of the conference. A suggested objective of a measurement standards information center is to process information of this type in order to minimize the time lapse between the verbal presentation of the information and its appearance in written form.

Question. "The suggestions and proposals of the NCSL will require expensive financing. What are the possible or probable sources of such financing?"

The most obvious source is membership dues or contributions. Various classes of membership could be set up with various membership fees, ranging up to possibly a few hundred dollars per year. Another source of funds would be fees for individual services rendered. The success of the program will depend, of course, on the value of services that are made available. The NCSL approach to date has been a parallel one, including some study of organizational matters but also an action program to develop the type of information that will be of value to standards laboratories.

Question. "To what extent do you consider that the work of NCSL overlaps the work of technical societies that have been working on measurements standards for sometime, and what active steps are being taken to prevent the activities of NCSL from duplicating unnecessarily the work of those societies?"

It should be recognized that NCSL looks at measurement standards and instrumentation problems from the point of view of standards laboratories and not primarily from the individual point of view. NCSL wants to maintain adequate liaison with other organizations and to request their assistance on such needs and problems as fall within their scope. A latter panel discussion will be devoted exclusively to liaison with other organizations.

Session 4. Corporate Measurement Standards Programs

Paper 4.1. Corporate Level Standards in a Decentralized Company

S. C. Richardson*

The corporate level measurement standards for the General Electric Company are maintained in the General Engineering Laboratory. Highly accurate electrical and physical standards are established and maintained in several fields of major interest to the Company. The prime purpose of these standards is to provide a reference for the certification testing of the standards of the operating departments. An associated function of this standards operation is the Interlaboratory Comparison of Calibrating Capabilities. This is a measurement agreement comparison program among instrumentation service components to constantly improve the accuracy of calibration throughout the Company. The paper describes the responsibilities and relationships between this corporate level standards operation and the standards and calibration components of the operating departments.

1. Introduction

A corporate-level standards operation and the standards operation of a product department or division have an important factor in common. Either type of calibration laboratory can be created to satisfy a new or newly recognized need, or can be the result of a gradual evolution and development to meet constantly changing needs. Considerable experience has been obtained in establishing product-orientated laboratories to meet the needs of newly manufactured products. Much less experience has been obtained in creating corporate level standards laboratories. It is the purpose of this paper to describe the evolutional development and functional operation of an existing corporate-level laboratory. The technical operation of this standards component has been previously reported [Richardson, 1959].

In the General Electric Company the corporate-level standards operation is part of the General Engineering Laboratory. This laboratory, which is actually a group of four technology-orientated laboratories, has developed ever since its inception to meet the constantly changing needs of the company. Because of this, it is interesting to briefly review its history.

2. Historical Development

The Standardizing Laboratory was organized in 1895 and began operation the next year under the leadership of Dr. Lewis T. Robinson. By 1909 the laboratory consisted of a general instrument, oscillography, and iron testing group; a switchboard instrument group; an instrument transformer and power measurement group; and a service group for design, drafting, procurement, and payroll functions. Dr. Robinson served for many years on the International Electro-Technical Committee and was also active in the original organization of the American Standards Association. His interest in both manufacturing standards and measurement standards was instrumental in establishing early conformance to accepted instrument manufacturing standards and the conformance of all GE standards to national standards. Company standards have been traceable to national standards ever since. In fact, right from the beginning "there was a strict rule that no instrument was to be used for testing unless it had been carefully calibrated within a reasonable time prior to use," [Miller, 1953].

The standardizing and engineering effort of the laboratory continued to develop and expand, and several notable engineering contributions were

---

*General Electric Company, Schenectady, New York.

made in World War I. A major increase in engineering capability came in 1919 when the Consulting Engineering Laboratory, which had developed under Dr. Charles P. Steinmetz, merged with the Standardizing Laboratory. The combined organization was named the General Engineering Laboratory. In 1945 another Consulting Engineering Laboratory, which had grown and developed under Dr. E.F.W. Alexanderson, was also combined with the General Engineering Laboratory. This merger, further broadened the laboratory's engineering scope and capability.

During this period of engineering growth and merger the original standards and instrument calibration activity also grew to meet company needs. In addition, a number of standards were established for use in the various engineering sections of the laboratory. They benefited both engineering and instrumentation developments. In 1953 nearly all these standards were combined into one corporate-level standards group. This meant that the original standards effort in d-c electromotive force, resistance, and a-c at power frequencies, was considerably augmented. Standards added were frequency-time, capacitance, inductance, energy (watt-hours), temperature, mass, force, pressure, and volume.

Along with this growth in the electrical and physical standards operation there was a much larger growth in the service group responsible for instrument calibration and repair. In 1958 this instrument service group, which had grown to a personnel of 50, was transferred to the Service Shops Department and became a separate operation known as Schenectady Instrumentation Service. They provide service to both Company and commercial customers.

3. The Present Standards Operation

At the time of the consolidation of all standards into one group the objectives of the standards function were reviewed. The main change was that calibration work of a secondary or tertiary nature was no longer encouraged. The objectives are five in number:

1. To establish and maintain highly accurate reference standards in the technical fields of major interest to the entire company,
2. To test and certify the reference standards of the operating departments of the company,
3. To conduct the Interlaboratory Comparison of Calibration Capabilities,
4. To provide consulting service in the fields of measurement standards and high accuracy measurement.
5. To perform special highly accurate measurements within the technical scope of the operation.

It has been the policy to concentrate only in those fields of major interest to the entire company. This involves a selection of the particular technical areas, the degree of penetration, and the range and accuracy over which standardizing is to be provided. In technical areas of major interest to only a very few departments of the company, it has not proven feasible to establish standards at the corporate level. The establishment of such limited application standards results in a large ratio of maintenance time to actual use time. The selection of the technical areas in which standards are to be established is one that requires the very best judgment in order not to waste resources. Most of the engineering effort required to develop and establish standards and associated measurement procedures has been provided by company funds. The actual purchase of the standards is financed by laboratory funds.

In regard to the second objective, that of testing and certifying the reference standards of the operating departments, such service is supplied on a cost basis. The entire system is a voluntary one, and there is no compulsion on any company component to send its standards to this laboratory. A measure of the success of this voluntary system is that between 40 and 50 operating departments annually send their standards to this laboratory for test and certification. At first thought this voluntary system would seem to conflict with present day military contract requirements. In the decentralized operation of General Electric the responsibility, authority, and accountability for a product rests with the product department. The corporate level standards function does not have the authority to monitor or police a mandatory recall. In effect, a department imposes a self-monitoring mandatory "send-in" on their reference standards to fulfill contract requirements as part of their product responsibility. It is an option of the product department to send their standards to this corporate level laboratory, the National Bureau of Standards, or to some other qualified laboratory. There is, however, one exception to this: a mandatory recall system, affecting only a few operating departments, has been established for certain contracts requiring mandatory compliance for periodic certification tests. This was done with the concurrence of the affected departments.

The Interlaboratory Comparison of Calibration Capabilities provides a measure of the agreement of all the participating components at the calibration output level. It provides an overall comparison that includes not only the standards of the participating component which may have been separately tested and certified, but also the techniques, associated equipment, and personnel. It serves as a spot check or technical audit on the calibration capability throughout the company and maintains a very high degree of accuracy in this activity. It provides an effective self-monitoring influence on all participants, including the standards reference group. This was a company-sponsored program for many years, but is now supported by customer orders. Approximately 40 different company components participate annually. This work is more completely described in another paper being presented at this conference [Richardson and Barnhart, 1962].

Consulting in calibration and standardization work is provided for all company departments. Incidental consulting is supplied as a company service, but extended consulting is financed by a purchase order from the interested department. In addition, the Electrical and Physical Standards group sponsors a company-circulated newsletter.

The final objective, that of providing special measurements, is almost self-explanatory. These measurements are based on the skills and capabilities in place and are provided at cost for any company component that requests them.

4. Performance Measures

The lack of adequate measures of performance at the corporate laboratory level is one of the most serious handicaps to such laboratories. The very fact that product quality is not directly dependent on a laboratory removes one of the strongest supports for any laboratory operation. Being close to the product is an advantage that an operating department standards laboratory has over a corporate-level laboratory. In considering whether to establish a corporate-level standardizing laboratory, there is never any question as to the need for standards. The questions always arise as to the extent of the need, and the number of technologies, their ranges and accuracy, for which standards should be supplied. To determine the degree of need, it is necessary to have convenient measures of the contribution of the corporate-level standard laboratory to an industrial company. This is extremely important. A technical performance measure that can be effectively obtained is the degree of measurement agreement maintained with the National Bureau of Standards. The certifications and test reports on interlaboratory standards sent to the NBS provide performance measures on the internal corporate level measurement programs used to compare and extend the measurement units at the reference standard level. Another performance measure is the effectiveness and the extent to which corporate standards are used by a company's operating departments. This measure can be numerically expressed in the number of departments serviced and the number of standards certified. This measure is direct and readily understood. It is most effective as a measure when the service is provided voluntarily. A complete mandatory recall operation would make this measure almost meaningless.

The most desirable performance measure would be in dollars. Unfortunately, the direct economic saving is very difficult to determine. A corporate-level standardizing operation, if it provides good service to the operating departments, can save those departments the cost of duplicate standards in many operations. An operating department can usually arrange to part with a given reference standard for a week or two for certification test, but cannot do the same thing on a prolonged basis. Prompt certification test service results in a large investment savings.

Another very specific contribution of the corporate-level standards laboratory is in the area of information dissemination and consulting. This contribution is again difficult to express numerically.

There is a definite need for better performance measures, particularly for the corporate-level standards laboratory. An excerpt from Lord Kelvin's famous measurement statement is appropriate "...when you cannot express it in numbers, your knowledge is of a meagre and unsatisfactory kind; .... ." It is ironical that a standards laboratory, whose very business is accurate measurement, should have insufficient quantitative means of measuring its own performance and contribution.

5. References

[1] Miller, J. A., Workshop of Engineers, first ed., p. 4 (The Maqua Co., Schenectady, N. Y., 1953).

[2] Richardson, S. C., Reference Level Standards in a Large Electrical Manufacturing Company, Instr. Soc. Am., Conf. Preprint No. 9-59, 14th Annual Instrument Automation Conf., Chicago, Ill. (Sept. 1959).

[3] Richardson, S. C., and H. D. Barnhart, Increased Confidence in Calibration Capability Through Interlaboratory Comparisons, 1962 Standards Laboratory Conference, Boulder, Colo. (Aug., 1962).

Session 4. Corporate Measurement Standards Programs

Paper 4.2. Administrative and Supervisory Concepts of the Measurement Standards Program at Battelle Memorial Institute

O. L. Linebrink*

A brief description of the services, the staff, and the equipment involved in the Measurement Standards Program at Battelle are given as background. Concepts of purpose governing selection of standards and associated equipment, calibration policies and procedures, liaison with project planners, and evaluation of services rendered, are discussed. Various approaches to the analysis of project accuracy requirements are presented along with the functioning of a center for instrumentation information and its communication channels.

1. Description of Battelle and its Instrument Laboratory

A brief description of the Battelle organization is necessary to fully understand the problems involved in its measurement standards program. Battelle is a not-for-profit endowed organization, dedicated to the advancement of science through the conduct and encouragement of scientific research. Originating in Columbus, Ohio, it has laboratories at Frankfurt, Germany, and Geneva, Switzerland, and a research station in North Florida for studies in deterioration of materials.

The Columbus Laboratory has over 2200 employees doing contract research at an annual rate of about 26 million dollars per year. It started in 1929 with a staff of approximately thirty people. Its more than six hundred contractors, or sponsors, range in size from the U. S. Government through large corporations, groups of corporations and companies, individual companies, and private individuals. It operates on a cost incurred basis. Its staff of research workers is organized into seven departments representing chemistry, chemical enginnering, engineering physics, mechanical engineering, metallurgy, physics, economics, and information research. With more than 50 divisions and numerous additional research teams, the scope of research covers research in 64 technical areas. The experimental research of course depends upon measurement for much of its data. There are an estimated ten thousand measuring devices used. Only about half of these supply experimental data. The rest are used for construction and control indications.

The Instrument Laboratory, while a part of the Physics Department, functions as a service group to the entire Institute. It originated about 25years ago as a common storage place for instruments and gradually included maintenance and calibration. At the present time over 200 standards and the facilities, and personnel trained in their use, are available. Six engineers, eighteen technicians, and two clerks are involved in the total instrument laboratory operation. This staff averages better than 10 years' experience at Battelle. About 20 percent of this effort is in the measurement standards category and the remaining 80 percent is about equally divided between maintenance and application.

While the original objective of the service was to provide more efficient availability and use of equipment, the confidence factor in the accuracy of research data gradually became more prominent. Instrument errors were guarded against by periodic comparisions with suitable reference standards. As experimental research demanded more and more accuracy of measurement, the requirements placed upon the standards laboratory equipment and personnel increased. One quarter percent reference standard meters had to be replaced with one tenth percent standards. One hundredth of one percent potentiometers, formerly standards laboratory equipment, became research laboratory working tools. Now ratio sets and ratio transformers with accuracies of the order of one part per million are required to adequately back up one phase of the measurements provided.

---

*Battelle Memorial Institute, Columbus, Ohio.

Automation in the form of automatic recorders, programers, proportional controls, multichannel pulse height analyzers, digital voltmeters and counters, rate meters, timers, etc., all require special equipment, traceable standards, and trained personnel to make their potential accuracies applicable to the research data.

2. Measurement Accuracy Responsibility

While the specific responsibility for the quality of the research data rests with the individual chiefs of the 50 some research divisions, the instrument laboratory group provides the central standards laboratory facilities and promotes new calibration services as an adjunct to quality control of data. It cooperates with the research staff on special measurement problems and, because of its familiarity with the calibration techniques and the accuracies involved, can thus secure better accuracies and more effective use of available instruments. Many measurements and the instruments used in making them are common in the needs of the various research divisions. Thus the instrument laboratory and its staff become the communication center on methods of measurement, availability of equipment at Battelle, sources and availability of equipment and services, from instrument manufacturers, NBS services, and traceability requirements. An important part of this communications center is a filing system consisting of manufacturer's literature, technical and measurement standards articles, operation and maintenance manuals, instrument calibration records, supply room and instrument loan, and location records.

This description may sound like a utopia, but there are never many dull moments for the laboratory supervisory staff. Our philosophy is that every measurement has an error. It is our business to know the magnitude of that error. Technical problems are not always solved, equipment needs are not always met, staff members need training on new equipment, accidents and malfunctions occur, trouble shooting is requested at unanticipated times and places, and last but not least cost of operations problems do occur.

3. Loan Pool Operations

Let us take a brief look at the administrative side of this service. Battelle's research projects are performed on a cost-incurred basis for the sponsors. This, of course, means that the accounting system for the instrument loan pool must be designed to charge projects only for specific service and equipment usage actually performed on the project. Also, service groups at Battelle are in general expected to be so organized that costs equal income. Two major sources of income to the instrument laboratory are use rate on loan items to projects (88%) and charges to projects for services rendered which do not involve loan equipment (12%). One of the accounting problems on the income side of the ledger is to properly assess the miscellaneous small services rendered, such as answering questions, spare parts and supply room service, supplying operation and specification information, minor repair or testing service, location and transportation of equipment, etc. At the moment we are taking a closer look at this problem.

Staff time accounts for a major portion (70%) of our expense. Depreciation, another name for installment payments on capital equipment purchases, is the next largest item (16%). Small, noncapital purchases and outside repair and calibration services are relatively small (7%). The department and division offices assess an operations cost (5%). The remaining cost (2%) consists of travel, technical services secured from other departments, and miscellaneous items.

Purchase rate of new instruments for the loan pool has over the years about equaled the depreciation charges, however, this is not the deciding factor when purchasing a new item. Versatile instruments which meet a variety of divisional needs, make the most suitable loan instruments. To evaluate the present and anticipate the future needs, one needs to be a combination of a Sherlock Holmes and a fortune teller. The first bases conclusions on facts, the second on intuitive and educated guesses. Instruments which are an integral part of a system which is permanently located or has full time use in a research division are usually purchased by that division. Likewise, instruments, which are so special that they meet the needs of only a single project, may be purchased by the division or department involved or in some cases by the specific project.

The main reason for some instruments decreasing in amount of use is the changing nature of the experimental work. It is seldom caused by the obsolescence or wearing out of the item. Some potentiometers and recorders are still in use after more than thirty years.

The problems of management in the purchase, maintenance, and calibration of instruments have become more pronounced with the current trend for the departments to purchase more of their own instruments. Sometimes these instruments are parts of permanently located systems. Others are not recognized as instruments on purchase orders on property control records. Consequently some overlapping and duplication of departmental "loan pools" and equipment have developed as new technical areas and divisional structures emerge. Production plant type regulations and policies, issued from top management, do not seem to be appropriate solutions. The interrelation as well as the isolation of research groups make a simple solution difficult. The solution to this situation is through a better mutual understanding of the needs and objectives.

4. Calibration Control

While the calibration and service of these divisional owned items are at the discretion of the division chiefs, the instrument laboratory staff suggests and urges, that those instruments from which data are taken, be included on the lists for periodic calibration. They have the cooperation of the purchasing department in reviewing many of the orders for instruments, for unnecessary duplication or conflicting interests with the loan pool. The instrument laboratory is alerted by the receiving room on the arrival of new instruments, so arrangements for their calibration can be made.

Of course complete control of the calibration of loan pool items and instrument laboratory standards rests with the instrument laboratory staff. To keep the rest of the data producing instruments adequately calibrated is another perpetual problem. "The instrument hasn't been used, since it was last calibrated." "We don't use it for accurate data." "Our work is slack, so we don't have the money to pay for calibrations now." "We can check it against other instruments ourselves." These are typical of the statements made concerning divisional equipment by divisional staff members. Divisional reorganizations, shifting of research groups and equipment, and staff turnover further complicate the calibration effort. One possibility being considered is that all experimental data records include the identity and last calibration date of the instrument from which it was taken, and that all calibrations be recorded with traceability shown.

5. Measurement Training

Acquaintance of, old as well as new, staff members, with the Instrument Laboratory's calibration program and services, has been largely left up to the instrument laboratory staff, through direct contact with the users of the instruments. Well satisfied customers help spread the word. Group discussions, both formal and informal, are occasionally arranged. Efforts are made by the instrumentation staff to keep informed of new technical areas of research and to anticipate the instrumentation involved. Currently lasers, fuel cells, vacuum, and biophysics are receiving special attention.

The selection and subsequent development of staff members for the calibration and standards laboratory have always been problems. Completely trained individuals, ready to do the tasks involved, were not and still are not available for hire. Academic backgrounds in physics, and mechanical or electrical engineering have proved most valuable for the professional staff. Prior training and experience in radio, television, or electronics have been quite valuable to our technicians. However, the one common and, I believe most valuable, qualification, is a genuine interest in the physical hardware involved in making a measurement and in learning to understand how it works.

An active, but cautious, curiosity coupled with a respectable amount of self confidence and initiative is a big asset. Directions must be followed in detail, and yet many small as well as large decisions are required daily. Absolute honesty and recognition of ones own limitation are required. Supervision has the responsibility of learning new technical areas and leading staff members to become interested and competent in those needed at Battelle. Much use of descriptive literature, operation and maintenance instructions, as well as encouragement in technical society type activities is helpful. Association and work with the many Battelle scientists and engineers in the widely diversified technologies, through the common denominator of instrumentation and measurements, provides an experience which is unique. Those comparatively few staff members who have chosen to leave our laboratory group have invariably gone into positions of responsibility and leadership in other divisions at Battelle or with other companies engaged in research or production. Individual development on the job is not only an opportunity, it is a requirement.

Battelle is noted for its team work approach to research. A vast array of experiences and abilities may take part in the solution of a specific problem. The instrumentation and standards laboratory staff and its measuring equipment become an important part of many of the team approaches to the solution of research problems. In some respects it is comparable to the equipment manager and water boy on a football team. Frequently, however, it has the key which unlocks the door to research progress.

Session 4. Corporate Measurement Standards Programs

Paper 4.3. Corporate Measurements Standards Real and Abstract

L. R. Wallace*

The difficulties encountered in maintaining proper measurement agreement within a laboratory are well known by persons in the field of measurements. The problems multiply many times when this same measurement agreement must be maintained between laboratories widely separated geographically. This paper will present a method whereby corporate measurement agreement can be maintained between various divisions without recourse to one central corporate standards laboratory. To achieve this end, it is necessary to define the measurement standards function at the corporate level. It then remains to be decided to what extent measurement agreement is required. This having been accomplished, the techniques learned in the laboratory must be applied in such a way that the necessary agreement is achieved. Having achieved this agreement, a means must be established whereby continued agreement is maintained. A method presented utilizes the standards laboratory information center and planning functions.

1. Introduction

It is a generally accepted axiom that measurement standards are necessary in an industrial corporation if that corporation is to continue to function properly in all its phases. The measurements needs of our industries are varied and no one program can satisfy all these various needs. It will be the purpose of this paper to describe a method which we believe to be effective, in a divisionalized corporation, in maintaining adequate standards for the corporation without having a central corporate standards laboratory. There are four basic functions in this program--information center, corporate measurement standards, planning, and measurement agreement. They will be discussed in that order.

At the start of the program in 1959, the several divisions had their own separate standards laboratories. As can be expected, each was at a different level of advancement, and each had its own peculiar problems with which it was concerned. The first task to be accomplished was to find areas of common interest, determine which standards were applicable, and to get agreement within the corporation for these standards. To accomplish this mission, a corporate measurements standards task group was formed. Each division was asked to appoint a representative to this task group. It was apparent to this group at the very beginning that to be at all effective, the choice of standards had to be narrow. Therefore, the field was narrowed to those standards which were normally returned to the National Bureau of Standards (hereafter referred to as NBS) for periodic recertification.

At this point, a word of explanation might be in order. Each location participating in this program had three levels of requirements. The production and test departments were the ultimate users of equipment. This equipment was maintained accurate by Calibration Departments and by Tool & Gage Inspection Departments. The standards of the latter departments were certified by the Standards Laboratories. The standards used in these laboratories are the ones we are concerned with at this time.

2. Information Center

Initially, it was difficult to know just which of the many disciplines involved should be first considered for the establishing of corporate standards. It was here that the information center function was utilized. The standards laboratories in each location were encouraged to serve as local information centers for measurement. It was

---

*IBM Corporation, Data Systems Division, Kingston, New York.

felt that in this way a feedback loop would be established between the task group and the requirements in the various divisions of the corporation which would expose areas in need of immediate attention. The personnel of these laboratories were advised to actively seek out areas where they could assist in developing specific measurements techniques, advise on the procurement of instruments, and make measurements when specifications required the use of precise equipment and controlled environment. Thus, the first use of the corporate information center was to help implement the program which generated the information center. We were "boot-strapping" even before we had established corporate standards.

Divisionalization implies the manufacture of different products. These products in turn require different kinds of measurements. It is here that this program has an advantage. When a query pertaining to measurements cannot be answered locally, the total corporate experience can be brought to bear on the problem by presenting it to the task group. The group can individually or collectively search out the answers through knowledge of and access to many diversified corporate functions. If the required information is not forthcoming, they are empowered to seek outside assistance. At the same time, the group is made aware of a possible exposure in the corporate ability to measure adequately. An effective means of maintaining an adequate program is through the task group's periodic review and analysis of the local laboratory activity in providing information and guidance. Local questions about how to measure some phenomenon with increased accuracy or precision, requests for measurement of some parameter which requires the use of high level standards and requests for correlation of some measurement not previously made all serve to alert the task group to total corporate need in the field of measurements standards. This is a continuing activity.

3. Corporate Measurements Standards

Returning to the first use of the information center, we had determined from the results of surveys and reviews which phenomena should be considered first, and it was felt that the various levels of standards should be classified. Even today there is wide misunderstanding in the field concerning the meaning of such words as primary, prototype, secondary, working, interlab, and transfer standards. Letter designations were favored until such time that word definitions can be agreed upon.

By mutual agreement, it was decided that there would be three levels of standards designated "A", "B" and "C". "A" level standards were those standards representative of the highest level standards for the several laboratories, and therefore, for the corporation. "B" level standards were those standards which were certified by using "A" level standards. These "B" level standards were then used to certify any "C" level standards used in plant. Referring to the previously mentioned requirement levels, the standards laboratories kept "A" and "B" level standards. Calibration and Tool & Gage Inspection used "C" level standards. In some instances, there were requirements for "B" level standards in various engineering laboratories. All standards were referred back to appropriate standards laboratory reference. Not all levels of all standards would necessarily be available at all laboratories. Where two or more laboratories each maintained similar "A" level standards, the results of intercomparison provided the corporate standard for the associated phenomenon. Where only one laboratory had a particular "A" level standard, it became the corporate standard for that phenomenon. Such a designation system determines the requirements of the highest level standard on the basis of corporate need rather than on the thesis of buying the best that is available. Sometimes the best available is not good enough. More often, a little ingenuity enables one to obtain satisfactory results with equipment on hand.

We now have similar "A" level standards for various phenomena located in the several laboratories--these standards are real. We also have corporate standards for the same phenomena, but these are abstract. They are obtained through recognition of the limit of our ability to mutually measure a given phenomenon. This is accomplished through the corporate intercomparison program. Each laboratory is responsible for the maintenance of its own standards. Where it ceases to be self-sufficient through "boot-strapping" or other means, it must then use the facilities and services of NBS. However, certified standards of the "A" level category, periodically intercompared by the several laboratories, give assurance that all laboratories are in mutual agreement concerning not only the validity of their standards, but also the validity of their measurements technique. Therefore, we can obtain true measurement agreement which is the basis for a corporate standard; namely: this particular phenomena is measurable to "x" degree at any laboratory in the corporation. The abstract nature of this corporate standard does not permit direct comparison with a physical entity. It does, however, assure consistency within the corporation for each standard so intercompared. At the same time, it permits the greatest latitude of individual plant freedom of operation in keeping with the concept of divisionalization.

4. The Planning Function

Periodic intercomparison of "A" level standards provide suitable data for the maintenance of the corporate standard. When the task group has information which shows that the ratio of accuracy between the standard and the variable being measured is approaching an unacceptable figure, the problem is reviewed to determine which locations require the increased accuracy. If all locations suffer from this problem, steps are taken to upgrade all affected "A" level standards. If one or more, but a minority, of locations have the problem, a projection is made to see if it is just a matter of time before all locations will have to upgrade their standards, or if the requirements are localized. In the first instance, a program is outlined for the gradual upgrading of standards at all locations. In the second instance, a program is planned which will eventually provide for the corporate standard being the combined efforts of the locations affected. This assures the highest level corporate standard without undue duplication of effort.

This type of activity will obviously point up the need for improved standards. New product developments present a more subtle means of bringing about obsolescence of standards. Various groups in each location are responsible for reviewing the effects of introducing new products and for informing those concerned of the requirements of these products. An effort has been made to make each of these groups aware of the role of measurements standards so that the laboratories are advised of new requirements. However, since self-preservation is a fundamental law of nature, so it behooves us to solicit information by inquiring into those areas which have been the source of our most confounding problems. Experience has shown that we must actively seek information which will assist our program since we are the most informed as to what information we need.

Prototype standards are touched upon in the discussion of measurement agreement. The planning function of the task group regarding prototype standards is to predict, based upon new product activity, what standards might be needed wherein suitable standards are not now available. Assignments are made for the development of such standards.

Standardization between laboratories is limited to those areas wherein adverse effects are evidenced because of lack of standardization. Although each laboratory has somewhat different requirements at this time, specifications for environmental conditions have been accepted by all. Ultimately, each laboratory will meet the same specifications as a result of the task group's planning activity. Certain measurements procedures are standardized when these procedures have been found to give consistently better results. These procedures are not made mandatory, however, if suitable measurement agreement can be maintained through the use of other techniques. Acquisition of standards is the responsibility of each local laboratory. However, in the interest of maintaining corporate standards, periodic review of proposed acquisition comes under discussion so that all corporate measurements needs can be best served. Again, duplication can be avoided by understanding the whole corporate requirement and acting accordingly. If divisions are geographically far removed, some duplication is unavoidable, but at the same time the resulting intercomparisons are the basis for another abstract standard. When distance is not a problem, doubling up on standards provides greater return for the dollar spent.

5. Measurement Agreement

Measurement agreement must be maintained within a laboratory, between laboratories and with vendors. Compatibility within the laboratory is the responsibility of that laboratory. It must show traceability of measurements to appropriate national standards. In showing this traceability, it has been the philosophy to utilize the interrelationships of the various disciplines so that dependence upon NBS for certification is minimized.

For a measurement to be at all meaningful, it must convey quantitative values recognizable in today's technology and it must be reproducible within specified tolerances. This reproducibility should not be restricted to results obtained by utilizing specific pieces of equipment in the measuring process. The measurement agreement function is the most vital aspect of the whole measurements standards program and is predicated upon awareness of need. There are several needs that must be met. Specifications have been established for devices and it is necessary to verify that the devices meet these specifications. The nature of the work is such that the specifications have very likely been pushed to the limit of present ability to measure. Therefore, the first need is to acknowledge the limitations of the measurement in determining the worth of the device being measured. The second need is to be cognizant of the conditions of measurement which might adversely affect the device being measured. Typical of such conditions are overloading or overstressing and environmental changes. The third need is for the realization that the measurements made must be in agreement with similar measurements which have been or might be made on the same device.

To maintain compatibility between laboratories, an intercomparison program is used. The task group chooses first the phenomena which are most common to all divisions. They then determine the conditions under which the measurements should be made and specify which data must appear in the reported results of measurement. Suitable recognized standards for these phenomena are contributed by the participating laboratories. The standards are certified by NBS and are then circulated among the laboratories. Each laboratory in turn measures the standard in terms of its own standards and forwards the data to a central point. All results are plotted against normal variations expected for the particular standard. The spread of the plotted values and the variations from the certified value are an indication of corporate ability to measure that particular phenomenon with respect to national standards. These results are further analyzed. If the results of a particular location are found to be significantly removed from the average of the group, the methods used by this location are scrutinized. If the values are considerably better than the group, the techniques used to obtain these values are publicized so that good features may be incorporated by others. If, on the other hand, the values are considered out of acceptable limits, assistance is dispatched to that location to help bring the values into acceptable limits.

So far, measurement agreement has been maintained through use of commercially available standards which usually have sufficient history to minimize the problems of agreement caused by the device alone. In areas wherein suitable standards cannot be purchased, the problems are magnified many times. A standard in this category will be referred to as a prototype standard. When it is found necessary to obtain a prototype standard, each location involved in the associated field of measurement is called in for consultation. The specifications for the standard are reviewed and a suggested approach is determined. The development of the standard is assigned to one or more locations depending upon the skill and experience available which is most likely to produce the desired results. A standard is then developed along the lines which permit maximum reference to existing recognized standards. These standards are probably somewhat removed from the field involved but they serve as a reference point. When the prototype standards are considered satisfactory for use and appropriate re-certification techniques have been developed, the standards are then available for use and are treated in the same manner as other standards. This standard permits measurement agreement within the corporation, but does not in any way indicate the degree of compatibility with industry. This in itself is a separate topic and is being treated elsewhere in this conference, therefore, no further reference will be made here.

There is one other type of standard which is treated differently. When one location has a requirement for a standard which exceeds the capability of existing corporate standards, a review is made of the status of these standards. If it appears that no other location will soon need this higher level standard, the location having the need then procures the standard and becomes the certifying agency for the corporation for this phenomenon, other next lower level standards being referred to it. However, before the location is recognized as certifying agency, it must show through appropriate history that it can offer satisfactory certification. As other locations obtain similar higher level standards, the intercomparison program again is used to determine the corporate standard.

Vendor compatibility presents an entirely different type of problem. In this instance, there is no interchange of standards. Rather, compatibility exists or doesn't exist based upon acceptance or rejection of purchased parts. Such acceptance or rejection is affected by many things other than measurements standards. Except to act in an advisory capacity in the field of measurements, the standards laboratory activity is restricted to measuring submitted samples and issuing a report of the findings. Such activity is usually restricted to those measurements which by their very nature require the equipment, skill or environment usually associated with the laboratory. Activity of the task group in this area is to establish or recommend appropriate limits for measurement which in the opinion of this group can reasonably be expected with the present state of the art.

6. Summary

A need existed for measurement agreement within the corporation which had been divisionalized. Geographical locations prevented having a central corporate standards laboratory. A plan was evolved to assure corporate measurement agreement of the highest level standards used in the various divisions by utilizing an intercomparison program. To further the effectiveness of the total program, a corporate task group serves as a central source of information related to the field of measurements. This group also performs planning functions to guard against obsolescence of standards and to prepare for standards which future products will require. The results to date have been gratifying, but lest one be led to believe we have a panacea, it must be stated that we are not without problems. We believe that the program will become self-perpetuating, but we feel intuitively that sincere, personal interest along with technical capability is the true keystone to corporate measurement standards.

Session 4. Corporate Measurement Standards Programs

PANEL ON CORPORATE MEASUREMENT STANDARDS PROGRAMS

Moderator:
Lewis R. Wallace, IBM Corporation
Kingston, New York

Members:
Peter Joeschke, Autonetics
Downey, California

Orval L. Linebrink, Battelle Memorial Inst.
Columbus 1, Ohio

Sheldon C. Richardson, General Electric Co.
Schenectady, New York

Robert M. Shaw, Westinghouse Air Arm.
Baltimore, Maryland

Summary

Question: "Assuming that various measurement standards activities are taking place in an organization, with perhaps standards laboratory functions in several different divisions, how can you determine when the activities should be consolidated into a corporate program?"

Mr. Joeschke:
Many factors have to be considered before any organization can make such a decision--factors such as products or similarity of products that the various divisions assemble and the requirements that a portion of them be produced in one division and assembled in another. Not to be overlooked is the problem of geographic location. It is very difficult to make a general recommendation.

Question: "How much measurement agreement is needed in various measurements which are made at various locations? I think the degree of measurement agreement required for a particular product is one of the most significant factors in determining whether a corporate level program is needed."

Mr. Shaw:
If you have any calibration system at all, you need a program to keep it up-to-date. If you are doing commercial work, it's academic as to how you accomplish this. If you are doing defense work, it is mandatory that you have a program. Now this program need not be your own. The question of mandatory recall brings out this point very well. Just about every defense agency has its own view on the subject. But one thing is common: you will have a program. If you do not have your own program, then you might have to follow theirs. The economics of the situation says it is much cheaper to have your own. How can you determine when a corporate program is needed? Your market, of course, is within the company. If there are not enough departments and enough standards, then a corporate effort is not economically justified, and each of the operating departments will equip itself as necessary to meet product requirements. The existence of a real need and the possibility of a real contribution to the company is the criterion as to whether initiative should be taken at the corporate level or whether you can operate satisfactorily on a department level.

Question: "How does the mandatory recall system work in a large company like General Electric?"

Mr. Richardson: General Electric is very much decentralized, and the responsibility, authority, and accountability rests with a product department. The corporate level of standards does not have the authority to establish a mandatory recall. Yet recall must exist, because it is a contract requirement. It is strictly the responsibility of the product operating department to meet the contract requirement, and he can establish it as a mandatory "sending" program. It is strictly his option as to whether he will send his standards to us, to the

National Bureau of Standards, or to some other qualified laboratory. There is no avoiding the fact that standards have to be maintained at top quality.

Question: "In IBM how is the task group set up and how does it operate in the corporate structure of the company?"

Mr. Wallace: The quality manager of each location determined the need for participation in this program. He appointed a representative from the standards laboratory at his location and, in most cases, this was the manager of the standards laboratory or a staff engineer of this particular standards laboratory. These members formed the task group and met together periodically. The reporting function that each member exercises is to his own immediate management. The task group also reports to management at corporate headquarters. This enables them to see what is going on and to make comments and suggestions. But they do not control the activities directly.

Question: "How does one get management support for a corporate program?"

Mr. Shaw:
There are various ways to get management support. If you have a contract that says you will follow some military specification, you are going to find out what this specification says and this you will try to sell to your manager. If he believes you and he thinks the success of the contract is dependent upon your getting the money you ask for, you will get the money. If he doesn't believe you, he may say, "Well, we have been building these things for years and all of a sudden you are telling me that we can't get along without your having a vast sum of money. I just don't quite see it". This is the hardest man there is to sell, and you may have to steer him just a little bit. You might show the relationship between contractual obligations and standards. You might translate this into the need for new equipment or a standards laboratory. If you are not close to the product and the requirements of the product which are specifically in the contract, you do not have that crutch, and it's a very good one. There is much more difficulty as you get more removed from the direct product. The portion of company funds spent in the standards corporate level operation is subject to frequent review.

Question: "How did the Autonetics operation achieve management support."

Mr. Joeschke: We stress quality to management, and we buy only what we need. We do not establish a laboratory for the sake of having a showcase. If we can perform the job with a little less, we do it. If the products require us to procure new equipment, we are able to substantiate it to management even if we have to go through a screening process. After we have established such a process, it is not very difficult to convince management.

Mr. Linebrink: We have good communications between the standards and measuring group and the actual research staff, and with top management. Our technical director founded the instrument laboratory years ago. When he is reviewing a report he frequently calls me before it is released and says, "(So and so) makes a statement in here as to the accuracy of the measurements. Can we back it up?" And when he asks me this, I say, "What instruments did he use?" We usually discuss this, and quite frequently I am very familiar with the planning of the experiment, the equipment which was used, and the personnel on the projects which resulted in that report, so that I can advise our technical director regarding the traceability or measurement agreement which we can expect. The confidence value is what he is looking for. How much confidence can we put in the research reports? And I think more and more if we can sell management on the fact that we can increase their confidence in our products, be they research or hardware, the question then of support becomes less severe.

Question: "Mr. Linebrink mentioned the program they have of loaning instruments. Have other corporation representatives investigated the leasing of instruments and calibration services? Have you examined the economics of these procedures?"

Mr. Linebrink: Our use rates at the present time are averaging about 6 percent of the purchase price per month. Thus we are paying for this equipment in something like a year and a half. That is only the depreciation cost. There is also maintenance cost, calibration cost, and others. At least in most cases, we are not providing service and know-how, but just the hardware. So don't compare the use rate of 6 percent which we have set up with the total cost of leasing an instrument.

Question: "Earlier there was a comment about never being associated with production. In our case, time and again Sampling Plan A and Sampling Plan B have to be carried out by the standards laboratory. If that isn't production, I don't know what it is. I wonder, to what extent do others do their production testing in the standards laboratory?"

Mr. Richardson: Since at least the early 1920's at General Electric, the standards operation and the general instrument repair and service operation have been separate but for many years they reported to the same manager. The bulk of the assistance to the production groups was given by the general instrument repair and service group. It is just that we were working at the higher accuracy level, not that we were in a protected atmosphere. So don't take anything I said to infer that we are on any elevated plane. We are very close to some of the production problems.

Mr. Joeschke: It doesn't matter so much what job you are doing as long as the company benefits from it. We have had occasions when production has come to us to get certain tasks performed in our laboratory, for instance, because the task was too small in size to obtain personnel or suitable equipment elsewhere. So actually we do production, too. But there is nothing wrong with helping anyone who needs help. Management expects some return for their investment. The better reputation you have, the better service you can perform. So thinking in terms of a laboratory, it helps definitely to keep the spirit of service. That is the mainstay. We serve.

Mr. Linebrink: Service is the keynote to our work. At Battelle, we have done some actual research project work, developing measuring devices to special needs. We definitely feel that if you can buy something, it is cheaper than trying to make it. But quite frequently such things are not on the market. We had a request from one of our research groups in chemistry which was developing some dental materials. This group wanted to test the strength of the materials used to hold false teeth in a person's mouth, and we developed a suitable measuring instrument. This is just an example of teamwork and reciprocity between the research staff throughout the Institute and the research staff of the instrument laboratory.

Mr. Shaw: In regard to the standards laboratory doing production work, it frequently happens that modifications to a product come along after you are in production. There are occasions when a measurement must be made on or near the production line and they do not have the facilities to do it. The prime consideration is always to meet and keep the production line rolling. If you don't have a standards laboratory or the function of a standards laboratory, you're not going to keep the line going.

Generally the economics of the problem shows that requiring the standards laboratory to do production work is justified only for a very short term operation. In the long run, it is very expensive. The equipment is necessarily much slower than production equipment. The gearing for the production is not there. The costing rate for the standards laboratory is considerably more than for the production line. When the foreman or the manager that's having trouble finds out how much it is costing him he will, in very short order, take the problem back where it belongs.

Session 5. Measurement Agreement Comparisons Among Standardizing Laboratories

Paper 5.1. Measurement Agreement Comparisons

W. J. Youden*

The best source of information on the measurement errors in comparisons is found in the records of the comparisons regularly carried out by a laboratory. This requires that some of the comparisons must be repeated, either directly or indirectly. An item, A, may be compared with a standard, S, and the comparison repeated. Generally it is better to plan the work so that the operator is not directly aware of how well his results check. If two items, A and B, are each compared with S, and then A and B compared directly, the additional measurement provides a check on the measurement process. Thus in addition to the direct comparison of A with S there is the indirect comparison obtained by adding $(A-B)$ to $(B-S)$. The sum of these two comparisons would check the direct result exactly if the measurements could be made without error. The discrepancy between $(A-S)$ and $(A-B)+(B-S)$ must arise from measurement error. Information collected over a sequence of such triads soon provides a sound basis for evaluating measurement error. This example and similar ones will be discussed in some detail in the paper.

1. Introduction

A calibrating laboratory must have in its possession appropriate standards with values certified by a competent authority. The calibrating laboratory must also possess adequate facilities for comparing its standards with items brought to it for calibration. The first thing the calibrating laboratory must attend to is to determine the accuracy of these comparisons. There are other problems such as the appropriate way to combine the comparison error with the uncertainty in the value assigned to the standard. This problem, incidentally, is only important when the comparison error is nearly as small as the uncertainty in the standard. This discussion is concerned with methods for ascertaining the accuracy of the comparisons, and also with getting the most information out of the measurements actually made.

2. Determination of the Accuracy of a Comparison Procedure

2.1. Two Independent Systems for Comparisons

It is not generally possible to attain absolute accuracy. Even if the calibrating laboratory has two similar certified standards and two completely independent assemblies for making comparisons, it is practically certain that, if enough items are calibrated with each of the two independent systems, a difference between the two systems can be demonstrated. This difference may be of negligible importance but once shown to exist, this difference is a component in the absolute error. Even when the calibrating laboratory shows this difference to be extremely small, there is the troublesome thought that the source certifying the two standards may have had some unknown error which was carried over into both certifications.

Such a series of duplicate tests with two independent systems on a succession of items furnishes the data for determining the accuracy of the comparison procedure.

---

*Consultant, Applied Mathematics Division, National Bureau of Standards, Washington, D.C.

Table 1. Data from two independent calibration systems

Item No.	A	B	C . .	N
System 1	a_1	b_1	c_1 . .	n_1
System 2	a_2	b_2	c_2 . .	n_2
Difference	D_1	D_2	D_3 . .	D_n

Examination of data tabulated as above should reveal whether the D's tend to be predominantly of one sign. The signs of the D's should, if the systems are equivalent, alternate in a random manner. The variance of the comparison process is estimated by calculating

$$s^2 = \frac{\Sigma D^2 - (\Sigma D)^2/n}{2\,(n-1)}$$

The square root of s^2 gives the standard deviation. This standard deviation (a measure of the precision) applies to any difference, Δ, found between a standard and a test item. It is this difference that applied to the certified value of the standard gives the value entered in table 1.

If the algebraic average for D is unacceptably large, this implies some persistent difference in the two systems. The obvious thing to do is to interchange the two standards with the two sets of comparison equipment. A further series of results will establish whether the discrepancy between the two systems arises from an inconsistency of the two standards or some lack of equivalence in the two sets of comparison equipment. Should the latter be the case, a suitable swapping back and forth of components of the systems will track down the source of inaccuracy in the comparison procedure [1].[1]

2.2 One System With One Standard

The usual technique for ascertaining the error in a comparison procedure is to repeat some of the measurements. This technique has the virtue of simplicity but it may not be the best way of obtaining data to determine the error in the comparison procedure. Direct repetition is vulnerable to repeating the same misreading of a scale. It is also vulnerable to "memory" or operator efforts to secure good checks. Few can resist the temptation, if a pair of results differs rather more than usual, to do one of two things--(a) To reject the pair of results and repeat the readings, or (b) To take a third reading and pair it with the closer of the first two readings. Many operators are unaware that if the average absolute difference between duplicate readings is R then about 11 percent of the individual differences legitimately exceed $2R$. If differences are rejected <u>solely</u> because they slightly exceed twice the average difference, the 'average' difference gradually becomes smaller. More stringent rejection will further reduce the average of the survivors. The logical end of this process is apparent reduction of the error to zero but at the price of rejecting all of the measurements. Another shortcoming of direct repetition is that there are alternatives that are slightly more efficient in estimating Δ, the difference between the standard and the item to be calibrated. More important, these alternatives reduce the number of times the standard is used and thus cut down on any wear or other consequences that follow from repeated use of a standard.

Quite commonly meter bar calibrations included not only comparisons of the standard with each bar but all possible comparisons among the bars in the group. Recently [2] the use of selected subsets of the pairings have been found satisfactory. On the other hand studies with standard cells tend to repetitive comparisons of a standard with the other cells and to make little if any use of intercomparisons among the cells. It seems likely that use would be found for schemes that replace most, if not all, of the repeat measurements by intercomparisons among a group of items only some of which are ever directly matched against the standard. This technique assumes that the test items are similar to and of comparable quality to the standard and also that the environmental control for the test items is equivalent to that maintained for the standard.

The principle of such schemes is shown by the example of comparing two items, A and B, with a standard S. We will suppose that the comparisons $(S-A)$ and $(S-B)$ are each repeated three times as is often done. Each set of three results provides an estimate of the variance with two degrees of freedom so the work provides a total of four degrees of freedom. A series of such sets of data will build up the number of degrees of freedom to give a better estimate of the variance. Note that the average of the three measurements of the difference between standard and test item has one third the variance of a single measurement.

A suggested scheme compares S with A, S with B, and A with B. Each comparison is repeated once. Observe that even if S and B were not directly compared, an estimate of $(S-B)$ is available by adding to $(S-A)$ the result for $(A-B)$. This information on $(S-B)$ can be averaged with the direct comparison of S and B. More weight is given the direct comparison. In this case the theory of least squares gives the direct comparison twice the weight of the indirect comparison. Denote $(S-A)$ by a, $(S-B)$ by b and $(A-B)$ by c. The weighted average for $(S-B)$ is given by $(2b+a+c)/3$. Similarly the weighted average for $(S-A)$ is given by $(2a+b-c)/3$. The variance for the average difference between standard and item when each of the three comparisons has been measured twice is again one third of the variance of a single measurement. Three degrees of freedom for error come from the three pairs and a fourth degree of freedom from the fact that $(S-A)+(A-B)+(B-S)$ should be zero in the absence of error of measurement. Consequently $(a+c-b)^2/3$ should be added to the sum of the squared differences of the duplicate readings. The square root of one fourth of this total gives s. If $(a+c-b)^2/3$ tends to be generally larger than the squared differences from duplicates, there is evidence of a certain amount of "forced" agreement between the duplicates. The scheme cuts the use of the standard by one third, retains the same variance for comparisons, and provides a check on the technique of measurement.

A scheme for three items (fig. 1, Scheme II) avoids the repetition of any measurement and cuts

---

[1] Figures in brackets indicate the literature references at the end of this paper.

the use of the standard in half. All possible six pairs of S, A, B, and C are compared. The average for $(S-A)$ is computed by combining five of the measured differences as follows

$$1/4\,[2(S-A)+(S-B)+(S-C)+(B-A)+(C-A)].$$

From symmetry, averages for all six comparisons are easily obtained. The six discrepancies between these calculated averages and the matching direct measurements reveal the measurement error. These discrepancies tend to be smaller than the differences between duplicates. The six discrepancies are squared. One third the sum of the six squares gives the variance of a single comparison. The variance of the average difference between standard and item is half that of a single comparison--just what the duplicate readings would give.

When there are four test items Scheme III, instead of duplicating the comparisons $(S-A)$, $(S-B)$, $(S-C)$, $(S-D)$, calls for comparisons $(A-B)$, $(B-C)$, $(C-D)$, and $(D-A)$. Now the calculated average for $(S-A)$ has a variance of 7/15 of a single comparison which is a small improvement over the 1/2 that simple duplication would give.

Scheme IV (fig. 1) reduces the use of the standard over Scheme III and provides for more information on some test items than on others. There are times when this discrimination among items is convenient. Scheme V reduces both the use of the standard and the number of measurements and hence reduces the degrees of freedom available for the variance estimate. In a continuing program this reduction in the amount of duplication may be acceptable if duplication is used largely to maintain a check on operations. Scheme V, interestingly enough, provides equal information on all four test items in spite of the corner position for the standard.

Schemes II, III, and VI are the first three of a series formed in a particular way. Beginning with Scheme III the comparison between standard and item has a smaller variance (about 7%) than straight duplicates would provide. The feedback through the comparison links brings about this improvement in efficiency.

Schemes VII, VIII, and IX show some additional patterns that may be extended to larger numbers of items. The tenth pattern illustrates a scheme making use of two standards. Clearly a wide variety of schemes can be devised. This permits the laboratory to select schemes appropriate for its particular program.

Two illustrative numerical examples are included. Formulas are not given for each scheme shown because they may be obtained from a statistician or a least square fit made to the data. A short cut for determining the weighting coefficients for the observed quantities is based upon an analogy with an electric circuit. The lines in the diagrams may be considered as one ohm resistances. If a potential is maintained between any two points the resulting equilibrium currents in the network give the relative weighing coefficients for the observations used to estimate the measurement comparison between the quantities represented by the two points to which the potential has been applied. Thus, in Scheme VII, if a potential of 3 v is applied between the standard and the midpoint of any side the current flow in the various resistances are exactly those shown in the first three lines of the illustrative example. A more detailed discussion is under preparation.

3. Interlaboratory Comparisons

It is common practice to send a "package" of several similar items on a circuit of several laboratories. The data should be examined to see if there is evidence that a particular laboratory tends to report consistently higher (or lower) values than the other participating laboratories.

One method of statistical analysis consists in taking the data for one of the items and assigning the rank of one to the laboratory with the highest value, the rank two to the laboratory with the next highest value, and so on. If there are L laboratories, the laboratory with the lowest value receives the rank L. This ranking procedure is carried out for each of the M items included in the package. A "score" for each laboratory is obtained by adding up the M ranks assigned to each laboratory. If a laboratory tends to get high values, its score will be low, but not lower than M. Low values lead to a high score with a maximum possible score of ML. If only random errors are responsible for the assigned ranks, the expected score is midway between M and ML or $L(M+1)/2$. Scores that depart sufficiently from the expected score constitute evidence of the presence of systematic errors. The attached table 2 shows scores which, if attained, constitute evidence of a systematic error. A detailed account of this new technique is available in Materials Research & Standards. [3]

Table 2

Let L laboratories test each of M materials. Assign ranks 1 to L for each material. Sum the ranks to get the score for each laboratory. The mean score is M(L+1)/2. The entries are lower and upper limits that are included in the approximate 5 percent critical region.

Approximate 5 percent two-tail limits for ranking scores

No. of Labs.	Number of materials												
	3	4	5	6	7	8	9	10	11	12	13	14	15
3		4 12	5 15	7 17	8 20	10 22	12 24	13 27	15 29	17 31	19 33	20 36	22 38
4		4 16	6 19	8 22	10 25	12 28	14 31	16 34	18 37	20 40	22 43	24 46	26 49
5		5 19	7 23	9 27	11 31	13 35	16 38	18 42	21 45	23 49	26 52	28 56	31 59
6	3 18	5 23	7 28	10 32	12 37	15 41	18 45	21 49	23 54	26 58	29 62	32 66	35 70
7	3 21	5 27	8 32	11 37	14 42	17 47	20 52	23 57	26 62	29 67	32 72	36 76	39 81
8	3 24	6 30	9 36	12 42	15 48	18 54	22 59	25 65	29 70	32 76	36 81	39 87	43 92
9	3 27	6 34	9 41	13 47	16 53	20 60	24 66	27 73	31 79	35 85	39 91	43 97	47 103
10	4 29	7 37	10 45	14 52	17 60	21 67	26 73	30 80	34 87	38 94	43 100	47 107	51 114
11	4 32	7 41	11 49	15 57	19 65	23 73	27 81	32 88	36 96	41 103	46 110	51 117	55 125
12	4 35	7 45	11 54	15 63	20 71	24 80	29 88	34 96	39 104	44 112	49 120	54 128	59 136
13	4 38	8 48	12 58	16 68	21 77	26 86	31 95	36 104	42 112	47 121	52 130	58 138	63 147
14	4 41	8 52	12 63	17 73	22 83	27 93	33 102	38 112	44 121	50 130	56 139	61 149	67 158
15	4 44	8 56	13 67	18 78	23 89	29 99	35 109	41 119	47 129	53 139	59 149	65 159	71 169

4. Summary

Calibration requires measuring the difference between a standard and a test item. Systematic errors can, with care, be practically eliminated from comparisons. Repeat measurements are generally used to estimate the precision of the comparisons. Repeat measurements may not be as independent as they should be. This paper lists various schemes that replace repeat determinations by comparisons among the test items. The advantages are (i) reduced use and wear on the standard; (ii) a more valid estimate of the precision; (iii) a slight improvement in the information obtained from a given number of measurements; and (iv) a flexible program adaptable to various programs.

A brief description of a new ranking procedure useful in interlaboratory tests is given together with a table.

5. References

[1] Youden, W. J., Experimental design and ASTM Committees, Mater. Res. Std. **1**, 862-867 (1961).

[2] Page, B. L., Calibration of meter line standards of length at the National Bureau of Standards, J. Research **54**, 1-14 (1955), RP2559.

[3] Youden, W. J., Ranking Laboratories by Round-Robin Tests, Mater. Res. Std. **3**, No. 1, 9-13 (1963).

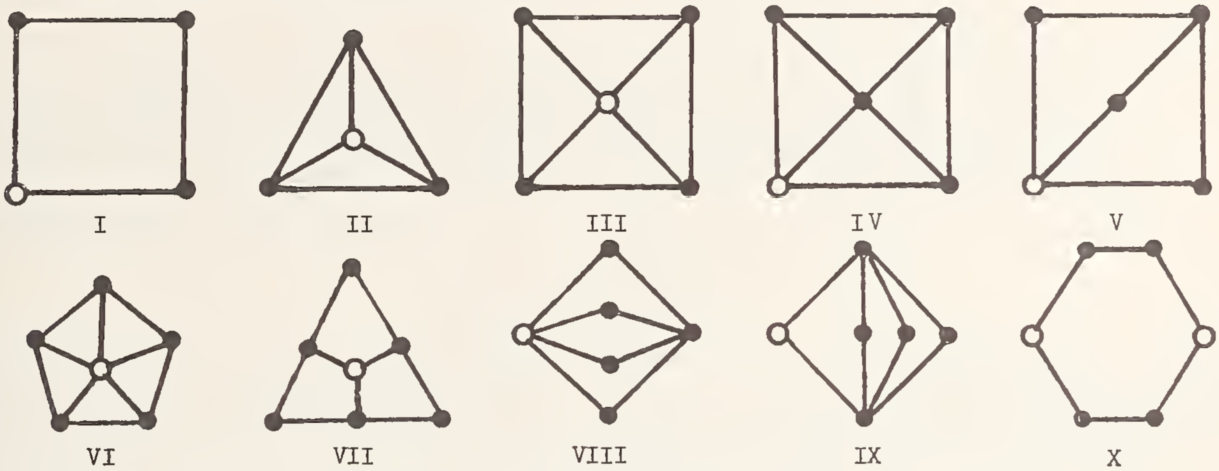


Figure 1. Calibration schemes. Circles identify the standard, solid dots represent test items, and connecting lines show the comparisons that are measured.

Illustrative examples:

Scheme I with data obtained in transposition of 10 gram weights.

Measured: S-A=-.011; S-B=.068; A-C=-.023; B-C=-.105 (mg.)

Calculated: S-A= $\frac{1}{4}$ [3(S-A)+(S-B)-(A-C)+(B-C)] = -.01175 ; S-B = .06875

A-C=-.02375; B-C = -.10425

S-C = $\frac{1}{2}$[(S-A)+(A-C)+(S-B)+(B-C)] = -.0355

Variance = Σ (diff. between measured and cal.)^2 = .00000225

s = .0015

Scheme VII using data taken with meter bars. See reference 2.

Nine pairings taken from a ten bar study using all 45 pairings.

Bar identifications: S=27; A=4; B=21; C=39; D=153R; E=752; F=814B

Pair observed	Measured value	Multiplying coefficients									Divide by	Cal. value	Obs.-cal.
		a	b	c	d	e	f	g	h	i			
S-A	a= 4.33	3	1	1	-1	1	0	0	1	-1	5	4.272	.058
S-B	b= -5.11	1	3	1	1	-1	-1	1	0	0	5	-5.090	-.020
S-C	c=177.13	1	1	3	0	0	1	-1	-1	1	5	177.168	-.038
A-D	d= 19.50	-2	2	0	7	3	-1	1	-1	1	10	19.469	.031
B-D	e= 28.80	2	-2	0	3	7	1	-1	1	-1	10	28.831	-.031
B-E	f=184.94	0	-2	2	-1	1	7	3	-1	1	10	184.929	.011
C-E	g= 2.66	0	2	-2	1	-1	3	7	1	-1	10	2.671	-.011
C-F	h= -6.96	2	0	-2	-1	1	-1	1	7	3	10	-6.933	-.027
A-F	i=165.99	-2	0	2	1	-1	1	-1	3	7	10	165.963	.027
S-D	(23.70)*	4	4	2	5	5	-1	1	1	-1	10	23.741	---
S-E	(179.80)*	2	4	4	1	-1	5	5	-1	1	10	179.839	---
S-F	(170.34)*	4	2	4	-1	1	1	-1	5	5	10	170.235	---

Σ(Obs.-cal.)^2 = 0.008830; Stand. Dev. = $\sqrt{.008830/3}$ = 0.054

*Measured by Page. Not used in these calculations to estimate S-D, S-E, and S-F.

NBS Miscellaneous Publication 248
Proceedings of the 1962 Standards Laboratory Conference

Session 5. Measurement Agreement Comparisons Among Standardizing Laboratories

Paper 5.2. Increased Confidence in Calibration Capability Through Interlaboratory Comparisons

S. C. Richardson* and H. D. Barnhart**

The object of the Interlaboratory Comparisons in General Electric, from its beginning nearly forty years ago, has been to improve the accuracy of instrument calibrations throughout the company. The specific purposes of this paper are to describe the operation of this measurement agreement comparison program, to outline the criteria that have been determined by experience, and to report the progress made in increasing confidence in the calibration capability through the company. Criteria are included on the method of operation, the comparison package, comparison techniques, and the reporting of comparison results.

1. Basic Description

The Interlaboratory Comparison originated in the mid 1920's and included six major plant areas. They were the West Lynn Works and River Works at Lynn, Massachusetts, the Erie Works, Fort Wayne Works, the Nela Park Laboratories in Cleveland, and the Main Plant at Schenectady. It was decided to use a group of 0.2 percent accuracy class portable electric instruments as the comparison group of instruments. Also included were resistors, current shunts, and unsaturated standard cells. Transportation was by automobile and two persons participated in the comparison, one each from the Meter and Instrument Department at West Lynn and the General Engineering Laboratory at Schenectady. Temperature measuring instruments and sensors were soon added to the complement of the comparison package. In the mid-thirties a set of proving rings were included so that the physical testing machines could be calibrated during the time the instrument calibrations were being made. For the past 10 or 15 years the comparison has included 40 to 45 company components. Approximately 25 of the laboratories participate in all phases of the comparison and the others in the calibration of physical testing machines only. Individual letter reports are issued to each participant in the electrical and temperature comparisons, and certificates are issued on the calibrations of all physical testing machines that are within ASTM specifications. At the end of the year an annual summary report is issued to all participating laboratories.

2. Method of Operation

Measurement agreement comparison programs are normally conducted by the use of common carriers for transportation of the comparison package. However, comparisons conducted in person have many advantages and are preferred whenever feasible. If a large percentage of the locations are less than 150 miles apart, a personally conducted comparison is best. Today, transportation is by station wagon with one General Engineering Laboratory man conducting the comparison with the participating group. A typical personally conducted comparison of two weeks would include about eight comparisons. The same number of comparisons using common carrier transportation would require 12 to 16 weeks, thousands of written words, and would be less effective. The present complete annual comparison takes approximately nine weeks. The entire comparison route is broken down into trips of one and two weeks duration. They are mostly one-day stops, although there are several two-day comparisons at plants with two separate participating groups. Experience has

---

*General Electric Company, Schenectady, New York.
**General Electric Company, Syracuse, New York.

shown that a two-week trip is about the maximum duration for efficient personnel operation and calibration reliability of the comparison package. In the typical two-week schedule, measurement comparisons are made during the day, and evening trips are made to the next location, distance permitting. The last trip of the year is always the one in a southerly direction from Schenectady. This is done so that, after the last comparison stop is made, the proving ring standards can be left at the National Bureau of Standards in Washington for test and certification in preparation for the next year. All trips are scheduled in advance with the laboratory managers who inform the government inspectors of the schedule so that they may witness the calibration of the physical testing machines.

3. The Comparison Package

In a measurement agreement program such as this, it is very desirable to have the comparison package contain items that represent the most common denominator for all participating groups. This is the reason that the portable electric instruments originally selected are to this day the heart of the comparison program. This electric instrument calibration is a normal function of each participating group. The groups range in size from 60 to 2 people. It is necessary also to have devices that cover an accuracy bracket of interest to all the participants. This is one reason precision resistors, inductors, capacitors, analytical weights, current shunts, and unsaturated standard cells have been included. A further requirement is that the comparison items be able to withstand the hazards of transportation. The temperature standards include several mercury thermometers, a base metal thermocouple, a noble metal thermocouple, and a platinum resistance thermometer. Five proving rings covering the range from 200 to 200,000 pounds are used for the testing of the physical testing machines and to measure the applied force on hardness testers. Class C dead weights are used to cover the range below 200 pounds. Several other comparison items have been included at various times but they have not stayed permanently in the package because of limited interest in them.

Great dependence is placed on the stability of the devices in the comparison package. It is therefore necessary to establish a historical record on them and use special care in their handling and transportation. Two identical sets of instruments are used. A spare set of the six instruments is maintained so that, if a given instrument becomes unstable or damaged, there is a suitable replacement available for the next trip. In this way the defective instrument can be repaired and the stability record re-established before it is put back into active service. It is necessary to establish deviation comparison limits on all the items in the comparison package. These limits are dependent on the stability of the particular comparison devices and/or the level that presents a challenge to the average calibration group.

4. The Comparison Techniques

All of the equipment in the comparison package is delivered to the participating laboratory for calibration by their normal procedures. They are instructed to calibrate these instruments using their regular facilities, techniques, and personnel. Experience has shown that while the instruments are not calibrated in a special way, it is human nature to calibrate more carefully. Instruments are set on zero and then energized for five minutes before calibration. These are the only prerequisites for calibration.

The same deviation tolerances are established for all indicating instrument calibration comparisons, even though the facilities vary considerably from location to location. An average deviation of 0.1 percent of full scale value has been established as the acceptable limit. This is a stringent requirement on instruments in the 0.2 percent class, but it has been proven satisfactory by experience. The actual measured values obtained in a comparison may be classed as a calibration, or as a cross-check. In most cases the measured values of the standard cells are obtained by a cross-check. However, the degree of agreement obtained by a cross-check may be as important to one group as that obtained by a calibration is to another group. Nearly all groups obtain measurement comparisons on all devices, but the ratio of those obtained by calibration to those obtained by cross-checks varies considerably. Some of the items in the comparison package are purposely selected or adjusted to have large scale corrections or values that deviate significantly from normal. Such an occasional large deviation is surprising and causes more concern than a more normal one. If the comparison of calibration results shows deviations larger than the established limits, an immediate effort is made to locate the trouble. This includes rechecking and the use of other methods and other equipment until the specific source of trouble is determined. During the time the comparison calibrations are made by personnel of the participating laboratory, the physical testing machines are calibrated against the proving ring standards by the person conducting the comparison. The physical testing machines are tested and force calculations are made immediately, so that it can be determined if rechecking, minor repair, or adjustment is necessary. Minor defects in the machine performance can normally be corrected. The physical testing machines are tested according to the American Society of Testing Materials specifications (ASTM E4-57-T).

5. Reporting the Comparison Results

The reporting method is semi-confidential, in that each participant is given a complete report on his comparison and only the summary report on the overall company comparison. This summary report, which is given at the end of the year, permits each participant to compare his performance with that of the average.

An unrestricted report is suitable for small groups and, in fact, was the form originally used. However, with a large number of participants the reporting of all comparison results to all participants becomes prohibitive and may on occasion be quite embarrassing. An informal oral report is made as soon as the calibrations are completed. The deviation for each scale point calibrated is reported in percent of full scale value, and the average deviation and maximum deviation are determined. The object is to get an immediate knowledge of comparison results to determine if rechecks are required. The measurement agreement is also determined on all other comparison items. At the end of the trip all calibration results and the deviations from Schenectady are supplied in a letter report. An excerpt from this letter report is shown in figure 4, which indicates data for other items contained with the interlaboratory comparison package. Recommendations are made for improvement if the comparison is not within acceptable deviation limits.

An immediate oral report is also made on the physical testing machines and hardness testers. At the end of the trip a formal certificate is issued on each of these machines that are within the ASTM specifications. After all comparison trips are completed, an annual summary report is prepared and issued to all participants. This summary report includes only the calibration deviations obtained on the electric indicating instruments. A separate histogram is compiled for each of the six instruments (figures 1 and 2) used in the intercomparison. This histogram lists all the deviations by all the laboratories in the terms of the reference group, which is the Electrical and Physical Standards operation in the General Engineering Laboratory. A histogram (fig. 3) is also compiled which represents the total of all the six instruments. This is plotted in a similar manner and represents approximately 2000 readings. Figure 5 is a plot of this data for the last twenty years, indicating the number of readings falling within the 0.05 percent full scale and 0.1 percent full scale. While this annual summary report and the seven histograms serve very well to present the entire year's comparison results, it is difficult for a particular group to compare their results against the company average. The ideal annual report would present similar information in such a manner that each participant could more readily compare his results against the average for all participating groups. Several methods of reporting are being investigated in order to further improve the benefit to all participants.

6. Results and Accomplishments

That this operation has been very successful in obtaining increased confidence in the calibration capability is evident in results obtained by an actual measure of the deviation improvements and by the acceptance the intercomparison has received by all the participating groups. It serves as a sampling check at the output level of the instrument calibration groups and provides the only independent check on operations that most participants receive. It serves as a technical audit on all participants, including the reference group. The histograms must peak very close to the zero deviation line when there are so many participants involved. It would be very embarrassing to the reference group to have these histograms peak at any significant deviation from the zero line. In this self auditing capacity, the comparison serves very well to improve and maintain the accuracy level. A subsidiary benefit of this comparison program is the improvement in the liaison between the company participating components. There is a large beneficial interchange of calibration experience and problem solving.

On figure 6, two sets of histograms indicate two very useful functions of interlaboratory comparison. Figure 6a contains data on measurement of unsaturated standard cells. The point of interest here is the deviation from zero in the plus direction as compared to Schenectady readings. A more complete analysis of this information indicates that these standard cells are compared against the participants own standard cells, which were certified on the average six months previous to the interlaboratory comparison. Therefore, there should be a slight offset in this histogram in the plus direction. This offset repeats each time that sufficient data has been available for good comparison. The histogram on figure 6b illustrates another form of data which a good interlaboratory comparison can produce. For instance, in obtaining resistance measurements, many levels of accuracy throughout the company are required to meet the requirements of the numerous laboratories and test areas. Therefore, the data shown here indicates that calibration laboratories with different degrees of accuracy requirements are fully serviced by such a program. It also indicates that such a program gets through to the needs of production line measuring equipment requirements.

The figure 7 histogram represents a certification record deviation summary for all physical testing machines which were tested and certified during the 1959 interlaboratory comparison. Here no data is reported beyond the one percent limits which are allowed by ASTM E4-57-T specifications. Physical testing machines that do not meet these deviation limits are not certified.

Many of the General Electric Company departments have no military calibration requirements; therefore, the yearly interlaboratory comparison program is welcomed as the chief source for meeting the commercial requirements of traceability or uniform measurements. Even though the interlaboratory comparison is performed by company personnel, a high level of respect is shown this program by all commercial and military departments. There is the normal extra effort by each organization previous to the scheduled date. In general, many of the department laboratories plan their own yearly cycle of self-calibration and cross-checking near the time of the interlaboratory calibration.

The results of many years of experience have yielded the following sequence of operations that best meet interlaboratory comparison program objectives:

1. Scheduling by Schenectady of the arrival date into your plant.
2. Arrival and unloading of comparison instruments.
3. Calibration by participant plant personnel of the interlaboratory comparison instruments.
4. Certification of physical testing machines using proving ring standards.
5. Immediate determination of accuracy of comparison (oral or written).
6. Reloading of the instruments into the station wagon for transportation to the next laboratory.
7. Feed-back of the semi-confidential report of the type shown in figure 4.
8. Compilation of the over-all company comparison records to supply each participating laboratory with total histograms showing the year of interlaboratory comparison (figures 1, 2, and 3).

Based on this type of program and feed-back, the participating laboratory can perform any corrective actions as deemed necessary after noting their deviations with respect to the Schenectady laboratory and the company as a whole.

One of the best measures of the success of any program is that which is obtained by dollars. Up until this year the interlaboratory comparison program has been a company sponsored project. It was decided that the comparison program should prove its worth by supporting its own cost. Consequently the participating groups for the year 1961 were informed that the 1962 comparisons would be performed only if authorized by a purchase order. Of the 42 comparisons made in 1961, 38 components authorized participation in the 1962 program.

Your authors represent the present administration and a participant with background or contact with several of the other laboratories. However, in writing of a program that has operated so successfully for many years, particular tribute is due to several past and present associates. Acknowledgement is made to Everett S. Lee, G. F. Gardner, and H. V. Miller for their contributions in establishing the basic pattern of operation, and to R. J. Pelletier for review and suggestions as seen from his viewpoint of present operator of this measurement agreement comparison program.

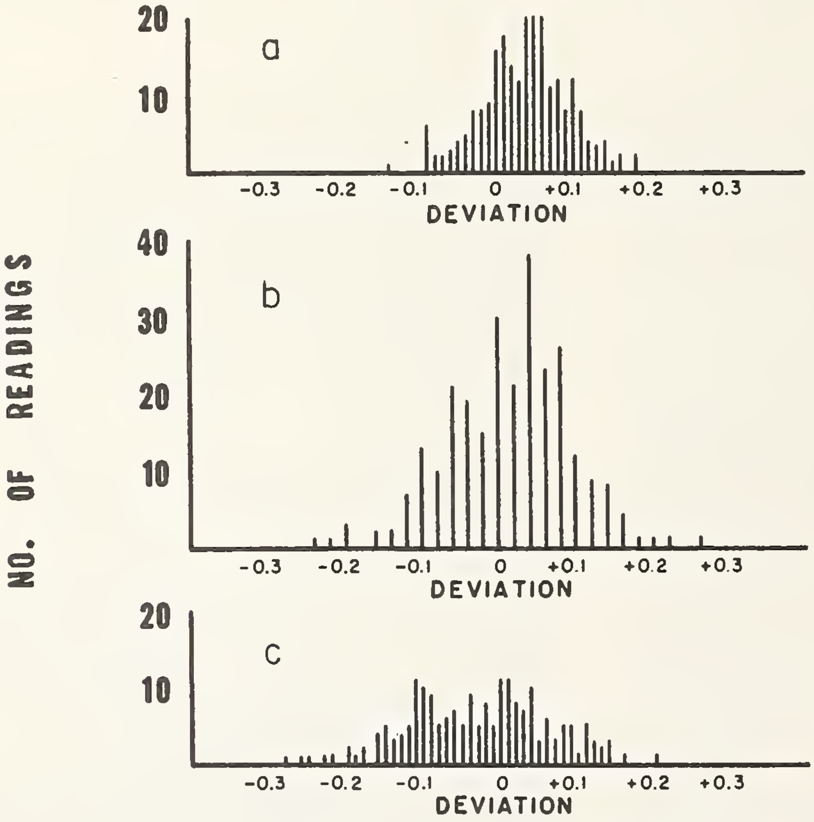


Figure 1. (a) 200 M.V. DC, (b) 5 Amperes AC, (c) 500 Watts AC.

+ INDICATES HIGHER READINGS THAN IN SCHENECTADY.
— INDICATES LOWER READINGS THAN IN SCHENECTADY.
DEVIATIONS ARE IN SCALE DIVISIONS.

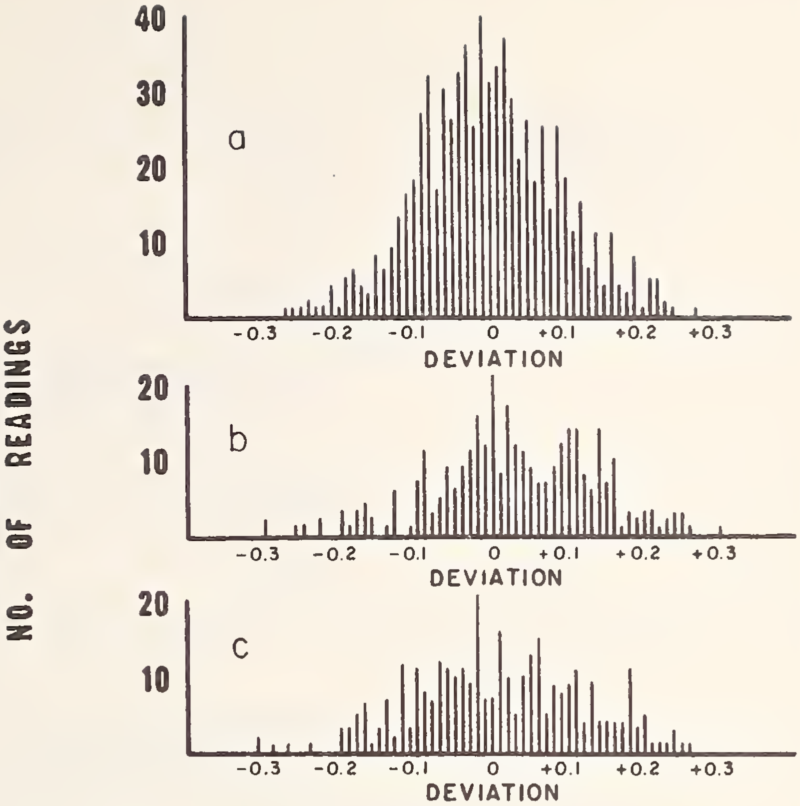
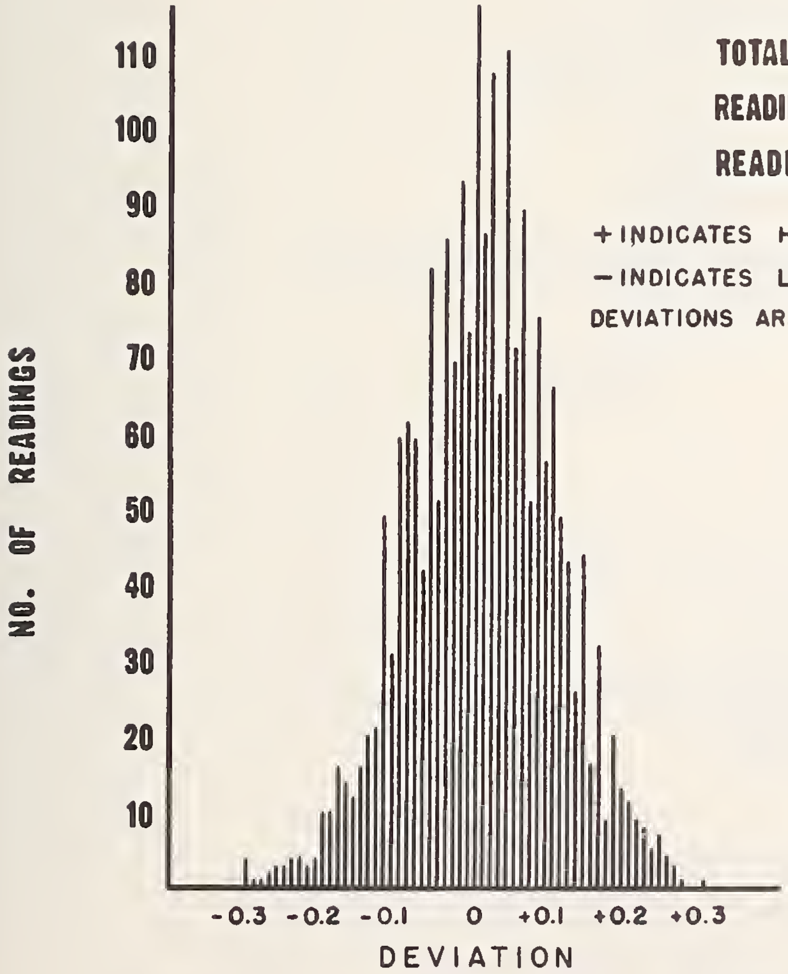


Figure 2. (a) 15/150 Volts DC, (b) 150 Volts AC, (c) 15 Volts AC.

+ INDICATES HIGHER READINGS THAN IN SCHENECTADY.
— INDICATES LOWER READINGS THAN IN SCHENECTADY.
DEVIATIONS ARE IN SCALE DIVISIONS.



TOTAL NO. OF READINGS = 2092
READINGS WITHIN 0.05 F.S. =56%
READINGS WITHIN 0.1 F.S. =86%

+ INDICATES HIGHER READINGS THAN IN SCHENECTADY.
— INDICATES LOWER READINGS THAN IN SCHENECTADY.
DEVIATIONS ARE IN SCALE DIVISIONS.

Figure 3. Total of all instruments.

Standard Cells
Value in Volts

Cell Number	G. E. Plant	Schenectady	% Deviation from Schenectady
11663-027332	1.01868	1.01867	+0.001
20803-048890	1.01948	1.01944	+0.004
20802-048891	1.01953	1.01952	+0.001

Average deviation from Schenectady in percent: +0.002

Precision Resistor #021202
Value in Ohms

Nominal Value	G. E. Plant	Schenectady	% Deviation from Schenectady
100	100.064	100.069	-0.005
1000	1000.37	1000.39	+0.004
10000	10039.8	10038.4	+0.009

Average deviation from Schenectady in percent: +0.003

60 Ampere Shunt #014275
Value in Ohms

Nominal Value	G. E. Plant	Schenectady
0.02	0.020012	0.0200114

Deviation from Schenectady in percent: +0.003

General Radio Capacitor - Serial No. 3085-048722
Value in Pico - Farads

G. E. Plant	Schenectady
1000.11	999.96

Deviation from Schenectady in percent: +0.015

Leeds & Northrup Inductor - Serial No. 118940-015965
Value in Millihenrys

G. E. Plant	Schenectady
100.16	100.314

Deviation from Schenectady in percent: -0.154

Figure 4. Other comparison measurements.

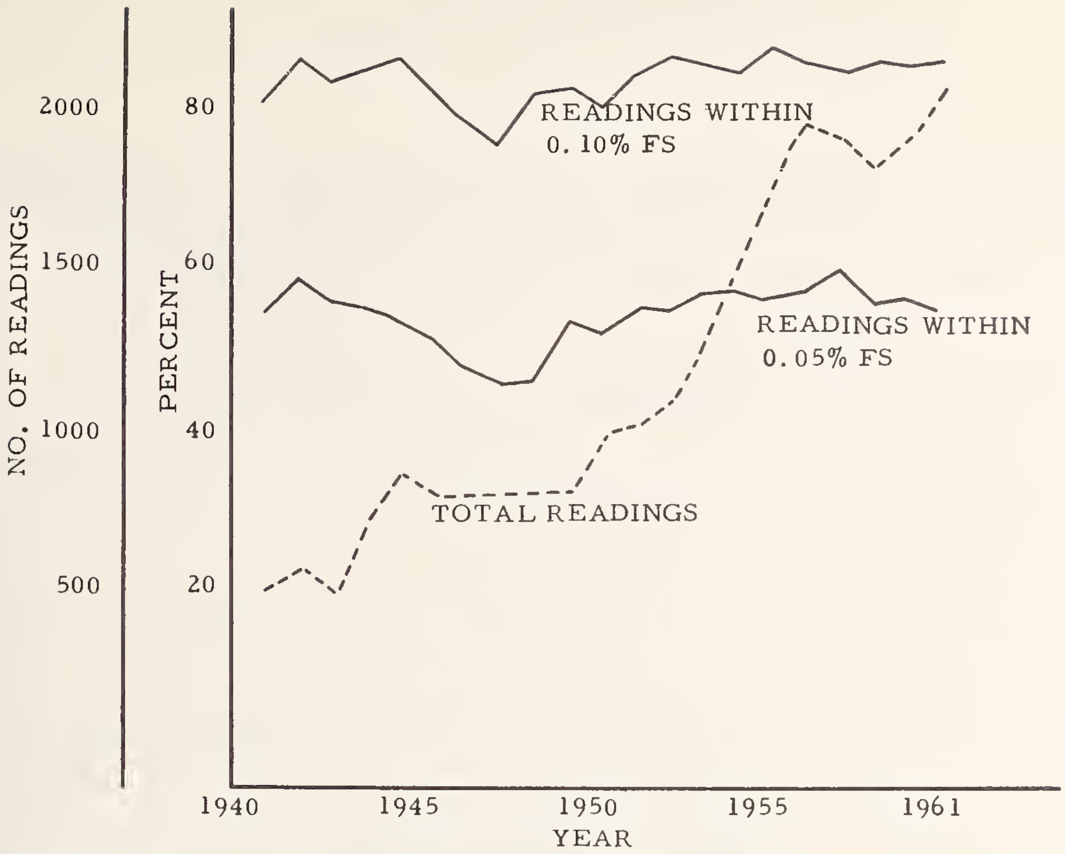
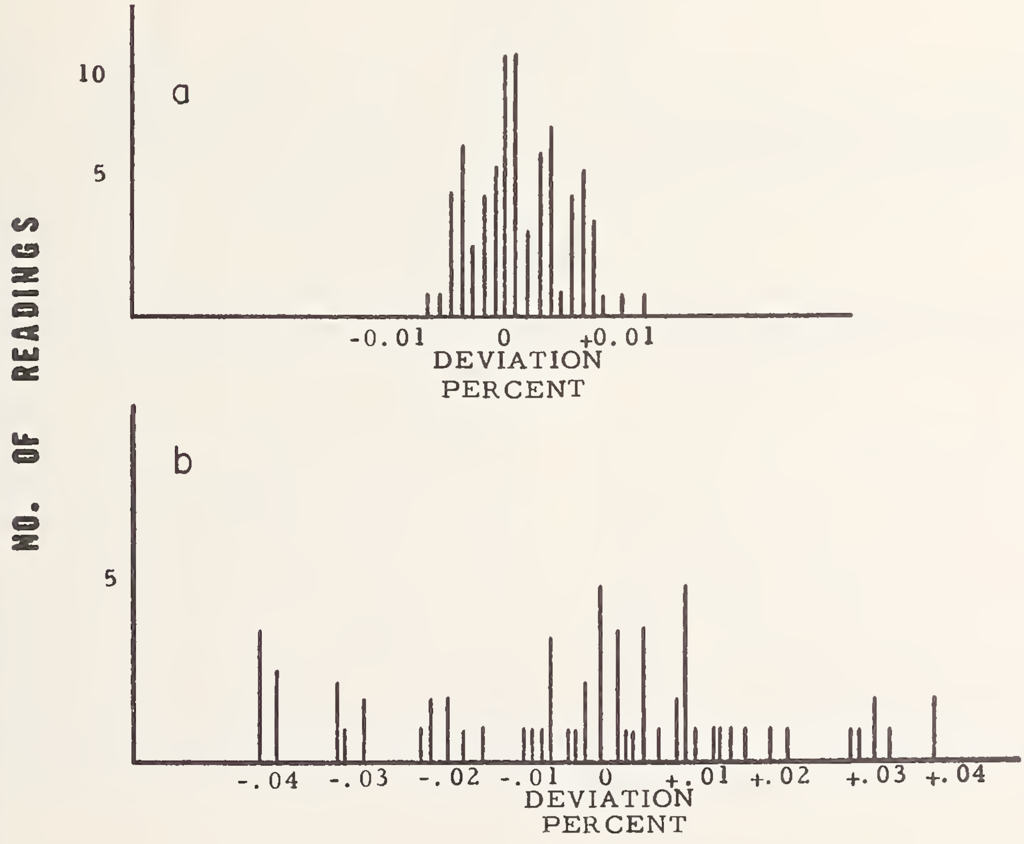


Figure 5. Summary of total instrument interlaboratory comparisons for last 20 years.



+ INDICATES HIGHER READINGS THAN IN SCHENECTADY.
– INDICATES LOWER READINGS THAN IN SCHENECTADY.

Figure 6. (a) Unsaturated standard cells and (b) Resistance.

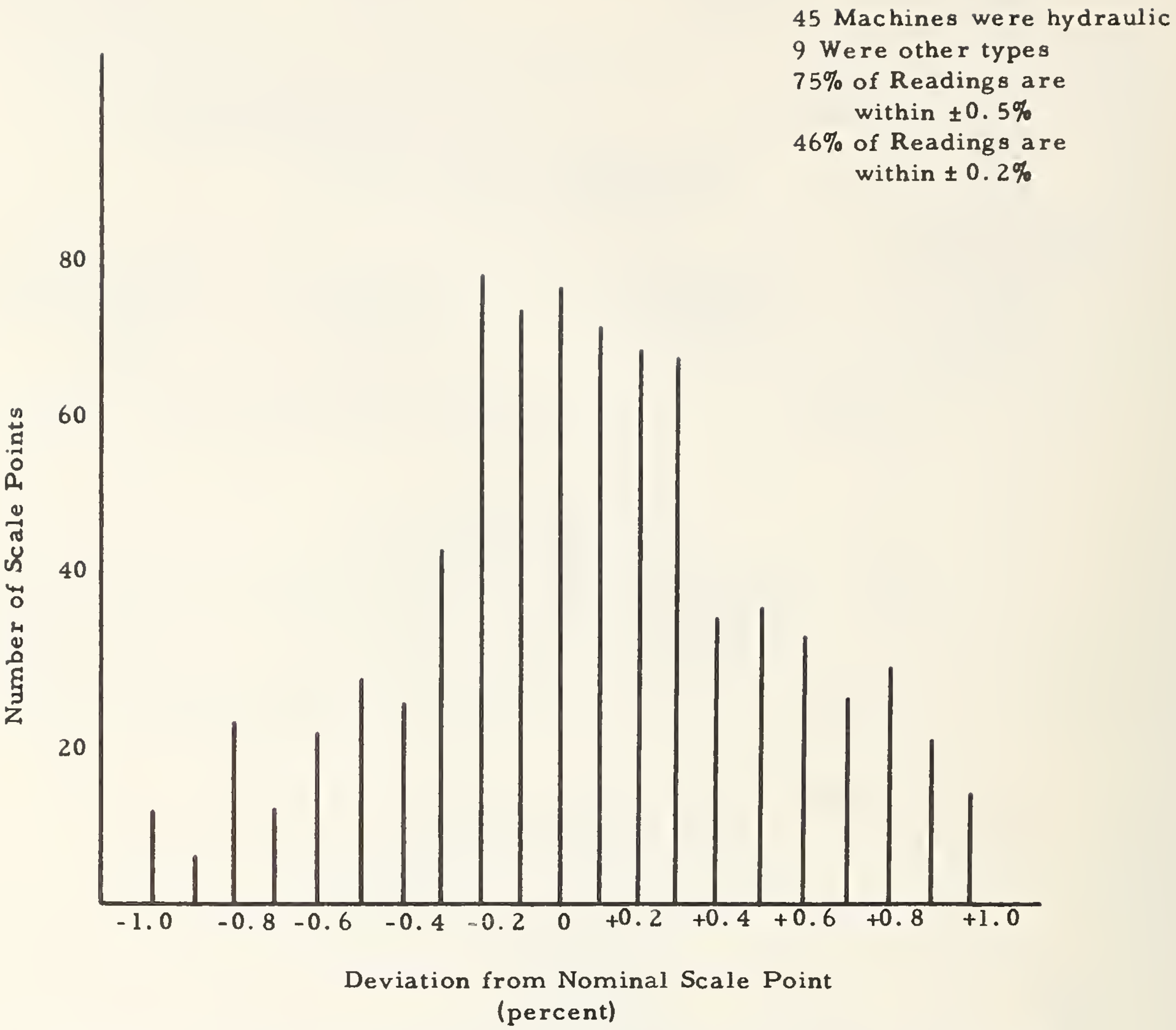


Figure 7. 54 physical testing machines standardized in 1959 at various general electric plants.

Session 5. Measurement Agreement Comparisons Among Standardizing Laboratories

PANEL ON MEASUREMENT AGREEMENT COMPARISONS AMONG STANDARDIZING LABORATORIES

Moderator:
Sheldon C. Richardson, General Electric Co.
Schenectady, New York

Members:
H. Curtis Biggs, Sandia Corporation
Albuquerque, New Mexico

Robert G. Davison, USAF Calibration Division
Newark, Ohio

Melvin L. Fruechtenicht, U.S. Army, Frankford Arsenal
Philadelphia, Pennsylvania

Jerry L. Hayes, Bureau of Naval Weapons
Pomona, California

Harvey W. Lance, National Bureau of Standards
Boulder, Colorado

Lloyd B. Wilson, Sperry Gyroscope Co.
Great Neck, New York

William J. Youden, National Bureau of Standards
Washington, D.C.

Summary

Question: "What are the armed services doing regarding measurement agreement comparisons?"

An interservice technical audit was started in 1960 at the Sandia Corporation operation of the Atomic Energy Commission. The Army, Navy, Air Force, and AEC-Sandia were the participants. The 1961 interservice technical audit included the same groups except that the Navy served as the host. The Army is the host in 1962, and the National Bureau of Standards has been included. The comparison items have included such items as coaxial microwave power mounts, X-band power meters, coaxial microwave attenuators, mercury-in-glass and platinum resistance thermometers, standard cells, standard resistors, four-terminal shunts, Type K potentiometers, Class S1 weights, gage blocks, and a barometer. The audit package is changed somewhat each year. Duplicate sets were included in 1962 to get more data for error analysis. Each of the three Services is also conducting intra-service or "in-house" comparison programs. These are conducted between the Type 1 and Type 2 laboratories. The comparison package is different in each case but contains items generally similar to those previously listed. The Air Force comparisons are conducted by a team of three persons from the headquarters laboratory. The Sandia Corporation has operated technical audit programs with the integrated contractors in the AEC program for a number of years.

Question: "Can there really be NBS traceability without NBS policing? Can it exist without NBS responsibility for instruments, standards, certification factors of accuracy, and basic standards?"

The word "traceability" was interpreted to mean that the users of standards at all levels refer the calibration of those standards back to the NBS or to standards that have been calibrated by the NBS. NBS traceability does in fact exist without NBS policing. Only in a rare case would someone knowingly claim something was derived from the national standards when it was not. Each laboratory must be basically responsible for its own work. There is no present system whereby one laboratory can be held responsible for some independent laboratory's work. That applies to the National Bureau of Standards. The NBS is not legally in a position to take responsibility for the work of others.

Question: "Is there in NCSL a single individual or a single group of people capable of telling an audience of laymen just how and why a system will fail to operate as a result of poor radioelectric measurements or lack of agreement?"

There are many individuals working in the field of standards who could illustrate and explain such possible system failures, but there is considerable difficulty in getting evidence of this sort. There is some inclination to shrink the responsibility for trying to point out such things, possibly due to embarrassing aspects of such situations, particularly after the fact. Simple illustrations can be given on how systems could fail because of inaccurate measurements. For example, radio receivers on aircraft often use push buttons to select fixed frequencies. If different frequency standards were used in Europe and in the United States and all the push buttons were adjusted in Europe, a serious situation would exist half way across the Atlantic Ocean, because the U.S. stations on a different frequency could not be received.

Question: "In diagramming the calibration of four instruments Dr. Youden suggested a square with the standard in the center and connected with the four corners (test instruments). If a tetrahedron were substituted for the square, would the accuracy of the operation be increased?"

The tetrahedron provides more information, and the accuracy should be increased. Two more comparisons would be required, but all possible pairs would be obtained. The tetrahedron represents an excellent design, but in this case it is more difficult to assign correct weighting factors to the measurements.

An electrical analogy may be used to illustrate the assignment of weighting factors. For instance, in figure 1, pattern III, of Dr. Youden's paper on "Measurement Agreement Comparisons", all the connecting lines could be considered as one-ohm resistances. If an EMF were applied between any two points in the diagram, the current flow would be along several paths from one point to the other chosen point. The relative currents in these paths are the multiplying factors for weighting. With sufficient experience, a person can look at these diagrams and assign quite accurate weighting factors by inspection.

Question: "Are the histogram results shown by General Electric due to random errors or systematic errors of the laboratories tested? If they are systematic, then why do you find the quadrature addition of the systematic uncertainty and random uncertainty objectionable?"

Both random and systematic errors were present in each of these laboratories. To show when quadrature addition is objectionable, consider this example. In a histogram made of the measurements or corrections on the three dozen or so prototype meter bars when they were compared with the international prototype, no one would maintain that the correction obtained on the NBS meter bar was the exact correction. It would be in error, and every time that meter bar was used to calibrate another, that same error would be included, unchanged and unaltered, until the NBS meter bar was taken back to Paris for another measurement comparison. These would be random errors from the viewpoint of the international comparison. However, from the viewpoint of the NBS they would be systematic errors because they are always included, in addition to the random error, in comparing the NBS meter bar (either directly or through a working standard) with meter bars sent to NBS for calibration.

The important point is that while a particular error from one viewpoint is random, the same error from another viewpoint is systematic. Quadrature addition is recommended, or not, according to the problem.

Question: "Can the technical audit type of program be more effective than the scheduled recall system?"

Both approaches are good, and both the technical audit and the recall system should be used. The emphasis depends on what a particular organization wants to achieve. A recall program is a part of the total measurements program, and the technical audit is a check on the entire measurements program. This is particularly so if the technical audit is extended to include the output end of the measurements program in which calibration is part of the product. A feature of the technical audit is that it is a check on the results, and the results are what is important. The technical audit will build confidence in a system. The laboratories that get good measurement agreement develop confidence. If the question were "Can a technical audit program supplant scheduled recall?," the answer would depend on the scope of the technical audit program. Good compatibility of measurements on 1-ohm Thomas type resistors might have but little meaning at the 10,000-ohm level. If the technical audit program could be extended to cover every known measurement at every measurement level, then there would be no need for scheduled recall. However, such an extensive technical audit would require so many people and would involve so much equipment that the effort would be greater than in a recall system. The technical audit program will establish confidence in the ability to maintain the unit of measurement but will not cover the entire range of parameters. The technical audit is basically a sampling technique, whereas the scheduled recall is 100 percent inspection.

Question: "What are the requirements for transporting standards and keeping accuracy disturbance to a minimum during travel? In your General Electric interlaboratory comparison how do you get the traveling standard cells to stabilize in such a short time?"

Transportation is provided by a company station wagon. All items are packaged in special carrying cases to withstand the normal road shocks and vibration of highway travel. Experience has shown that in the transportation of such equipment most damage occurs due to improper handling at each end of the trip. A personally conducted comparison eliminates this. Surprisingly, the transportation of unsaturated standard cells is not troublesome. Invariably there is time after an evening's run to let them stabilize overnight for the measurement comparison the next day. The reason three cells are carried is to provide an indication, by the measurement spread, of the disturbance to the cells. This spread normally does not exceed a few thousandths of a percent.

Question: "How should the audit package be designed? For instance, how do you select the standards to bring out the measurement capability you are interested in reviewing or surveying? Can this be treated as an experimental design problem?"

The Sandia-AEC annual technical audits are made to determine capability and to improve the measuring ability of the standards laboratories. For example, in 1955 dimensional measurement comparisons were made in the range of six to ten inches. The general ability of the laboratories in the AEC system was on the order of one to two ten-thousandths of an inch. However, some deviations were as much as one to two thousandths. The results in 1962 showed that the greatest disagreements were from one to two ten-thousandths, a considerable improvement. The best measurement agreement was five to ten microinches. The interest in technical audits is such that there are continual requests for more. A thirty-day time limit is put on the use of the audit package. The audit package must be simple, the measurement points must be well defined, and the stability must be good. Another important end result of a technical audit program is that it will reduce the need for a primary laboratory to serve as referee between or among secondary laboratories.

Question: "Do any groups limit the steps or the number of comparisons in the traceability exchange between primary reference standards and working level instruments in an effort to limit cumulative errors?"

There are limits on the number of steps in programs within the Navy. In-house programs are established with definite limits on the number of steps in comparisons between the different levels of calibration. It should not be implied that the military, because of their use of the phrase traceability, attach to it a limitation in the number of steps between the working level test equipment and the NBS. On a broad spectrum this is not so. An important factor is that the number of steps is sometimes counted by the number of laboratories between the level working on test equipment and the NBS. It is easy to overlook in organization charts the potentially significant difference in the way each laboratory operates. Four laboratories appearing in an organizational chart between working level and the NBS could have as many as seven echelons of calibration, depending upon how intercomparisons are made internally in the laboratories. For instance, if interlaboratory standards are used to check the reference standards, which are used to check the working standards, which are then used to check the standards of the next lower echelon, several echelons of intercomparison are involved. However, if the interlaboratory standard is called the reference standard and used as the working standard to calibrate the next echelon, the intercomparisons are reduced to one. The number of permissible steps is often based on each individual measurement technology. In the area of RF power measurements, there may be only one step allowable between the working level test equipment and the NBS. In the area of resistance measurements several steps can be used advantageously. Each service circumvents its various echelons as the state of the art demands. Internally the groups do limit, but the military has not extended this into contractual operations.

Question: "Since military contracts now require a calibration program with traceability, can contractor laboratories participate in the interlaboratory comparison tests with the service groups? Might participation become mandatory?"

It is a considerable task to conduct audits among the services in addition to the internal comparison programs. The amount of time and effort that is now available is used in the audit program among the services. It is not the present intent to make participation of contractor's laboratories mandatory in these programs. However, in cases of doubt, arrangements have been made with the standards laboratory of the contractor to make cross checks, and measuring devices have been brought to the military laboratories for comparison checks. Only to that extent have contractors participated. Contractor participation in general is not anticipated.

Question: "Audits represent a best behavior situation. Do any measures of actual operational accuracy exist?"

The Air Force and other activities, whether contractors or the services, use methods of auditing the actual results of the calibration laboratories and sampling the products that are measured. This is a normal part of most quality control programs.

Question: "We are not policemen. We are engineers. Why do we lose friends and alienate people by talking about policing?"

The National Conference of Standards Laboratories is specifically trying not to do any policing or to imply policing in any of its activities. The policing function in many individual cases is self-imposed. The fact that a measurement comparison is made does not mean that policing will be done. These comparisons can have a major and constructive effect in helping a standards laboratory to improve its operations. Laboratories can make their own technical audits. If some other group makes the technical audit, policing policy depends primarily on organizational relationships and on whether or not there is a desire to police.

Question: "How does the Sandia Corporation go about establishing true needs for higher accuracy and new measurements?"

The program for establishing true needs has five main aspects. These are to discuss projects with the designers, to estimate future measurement needs, to establish specific calibration programs to meet inspection programs, to develop new standards or different methods of calibrating standards in order to save manpower, and to endeavor to develop standards for other groups, so that they can make their special measurements and hence conserve the manpower of the standardization laboratory.

NBS Miscelloneous Publication 248
Proceedings of the 1962 Stondords Loborotory Conference

Session 6. Training of Measurement Personnel

Paper 6.1. Personnel Requirements for Calibrating Laboratories Military Practices

D. DeLauer*

This paper describes the personality characteristics, the types of skills, and where to locate measurement personnel.

1. Introduction

The need for highly qualified technical personnel in the military standards laboratory is the same as that for the industrial laboratory. However, certain differences exist in the overall operation of the military laboratories which necessitate a slightly different approach to the staffing and handling of the technician personnel.

The technician personnel are selected from within the Air Force and receive a high degree of training, but inasmuch as it is a military organization, there are many inherent differences in the operation of the laboratories. For instance, the fact that the purely military duties absorb so much time and come at such odd intervals means that laboratory supervision on occasion has no technical help at the bench. Or, in reverse, during an alert they may be on duty in the laboratory for a straight 24 hours. However, this is not the major problem which faces the Air Force Base Laboratory Calibration Program. This problem may be stated briefly.

In industry there is a tendency to specialize in a relatively narrow band of instruments, particularly as the level of measurement goes up.

In the Air Force, the Base Laboratory covers a wide range of precision measurements, yet it may be operated by as few as eight people because of the small overall workload. Can a technician do a good job producing measurements from DC- to X-band, from pressure transducers to platinum resistance thermometers, from surveyors levels to theodolites, aircraft scales to gage blocks, optical flats to autocollimators, geiger counters to digital voltmeters, and an endless line of special equipment? In addition, can the Air Force staff 178 of the laboratories in time to meet the missile and the ultrasonic aircraft schedule? The Air Force has said that it can be done.

The Air Force has evolved a system for selecting and evaluating its technical personnel at the base level which is felt to be extremely satisfactory for maintaining the quality of measurement in a high degree. How the Air Force has accomplished this for the Base Laboratory is the subject of this paper.

The Base Laboratories are in general about in a third echelon as referenced to the National Bureau of Standards. They are not physically placed in the third echelon in the Air Force Calibration Program; however, the types of equipment provisioned would normally be used in about that position.

The Air Force Base Laboratory has a big part in the responsibility for the accuracy of missiles, aircraft safety, and the degree of readiness of the Air Force armory.

---

*Air Force Calibration Detachment, Vandenberg Air Force Base, Calif.

The first question which may be asked is "What determines the broad base of measurements found at each Air Force Base?" which in turn places such a demanding requirement on the technical personnel.

Each base is practically self-sufficient. Each has a hospital, an airfield, including aircraft maintenance which may or may not be the primary mission, communications, civil engineering, safety, a munitions maintenance group, a weather group, a supply department, and a photographic group. At missile bases a separate squadron has responsibility for the weapon system operation and maintenance, and at each base the calibration laboratory is known as the Precision Measurement Equipment Laboratory (PMEL).

Each of these base activities has test equipment to aid in performance of its main function, and in all cases this test equipment is in a continual state of modification or updating. For example, the use of Wild T-3 Theodolites by Missile Squadrons has recently become common practice. The PMEL's supporting these squadrons have been set up to calibrate these theodolites--no easy task.

At many bases and always during the activation stages of a large scale program there are many contractors involved and many subsystems must be integrated and tested as a unit. It is understandable that instruments installed in these systems, if calibrated through a central base laboratory, will produce smoother integration and better results in future operations.

Thus, the base laboratory must be staffed with personnel capable of understanding and reproducing measurements of high accuracy to support these complex programs.

The broad base line of measurements is then a result of the self-sufficient nature of the base for military operation, the contractor requirement during the activation stage, and the complex nature of today's weapon systems.

To cope with the problem of staffing Base laboratories with workloads as varied as a large industrial laboratory, but with quantities which limit the number of personnel, the Air Force has decided to develop a well-rounded technician. It has taken time to accomplish, and with the past help of industry, the laboratories are keeping pace with the weapon system activation. The program is now functioning and the laboratories are producing good measurements.

Let us take a look at what is actually demanded of the Air Force technician:

1. Calibrate any instrument presented to the laboratory and repair most of them.
2. Act in the capacity of Quality Control, Production Control, and Supply Specialist.
3. Attend to his military duties, parades, physical education, war status duties, etc.
4. Develop good customer relationship.
5. Provide instrumentation information to all base activities.

To develop calibration personnel for the PMEL's, the basic requirements demanded may be stated as follows:

1. He must first be military career minded, considering the long training process in this field.
2. He must be in the very top group in general intelligence.
3. He must have previous test equipment experience.
4. He must have an aggressive nature in order to compete for promotions in PMEL work.

These are the characteristics and the requirements from which the Air Force began its search for PMEL personnel. It turned out that in almost all cases men were selected who had service amounting from 10-16 years, mostly in technical fields of radar, telemetry, aircraft maintenance, etc.

Today the selection process has crystallized in a firm structure and is the backbone of the calibration program. During any phase of the process, the technician may be "washed out" and returned to another career field. The officers and non-commissioned officers supervising the laboratories do not tolerate second best.

In selection of personnel to staff the laboratories, the Air Force has made almost no compromises. The selection is detailed and can be grouped into ten steps.

1. High School Graduate
2. 80 Per cent Aptitude Test
3. College Algebra
4. Two Years Test Equipment Experience
5. 2nd Enlistment
6. Pass Entrance Examination
7. Pass Basic Training
8. Pass O.J.T. Pre-Test
9. Pass Skill Level Test
10. Demonstrate Continuing Proficiency

To be promoted in this career field the technician must advance in the On-the-Job Training Program and pass tests covering all phases of calibration, including such things as detailed descriptions of test equipment circuitry, analysis of precision potentiometers and bridges, nuclear instrumentation, and optical calibration. After fulfilling these requirements, he must also be recommended by his superiors who judge him from a number of general considerations.

The PMEL program could not have worked by the selection and training of the technicians alone. Much credit has to be given to the Tech Order system for procedures and instruction manuals. This system, although having difficulty keeping up with the expanding equipment market, fills the gap in providing information on equipment which appears at the laboratory in such small quantity that familiarity is not developed.

Another reason why the program is working is that a high degree of interest is maintained by the technicians. This is primarily due to the fact that they are continually aware of the seriousness of their work and the fact that this work is their career. They are always faced with the possible explosion of a missile or crash of an aircraft which might be caused by a calibration error.

3. Conclusions

It has been stated that a very difficult situation existed in staffing such a large quantity of laboratories where only limited technical personnel were available.

Keeping pace with weapon system activation and producing measurements in support of a centralized program required a firm approach. This paper has given a generalized account of the situation and was not intended to imply that all problems are solved.

It is hoped that the original question has been answered. The technicians are producing good measurements over a wide range.

The general solution has been to select persons with the basic requirements, provide them with a long term training program, and provide incentive for promotion.

The system has been successful as evidenced by the fact that technicians have been transferred to bases with different missile systems and from missiles to aircraft without noticeable problems.

The program is working, and each year the professional nature of the laboratories is rapidly increasing as observed by us who visit them often; and, in spite of the promotion difficulties, when you walk into a PMEL, you will see that almost all are high-ranking Sergeants who go about their business with the air of professional standards men.

Session 6. Training of Measurement Personnel

Paper 6.2. Personnel Requirements for Calibrating Laboratories Industrial Practices

M. Hoskins*

This paper describes the personality characteristics, the types of skills, and where to locate measurement personnel.

A standards laboratory or metrology technician is harder to define than to find, but if you can't describe him you can't hire him.

The first step, in my opinion, is to know reasonably well the job that is to be done. For purposes of discussion I shall list a few general jobs.

1. The simple meter reading and recording type with no evaluation of findings necessary.
2. The simple set-up and meter reading with acceptance or rejection of performance the ultimate goal.
3. The more complicated set-up and determination of value through calibration or by rigid standards procedure.
4. The complete unknown evaluation requiring overall skills and multiple fact report.

There are several ways of selecting personnel for these jobs. It is possible to match each job to a person capable of that job only or to use a step-up system by training so that the technician advances through each type of job or a combination of personnel types.

The method chosen must be determined by the type and quantity of work in each standards laboratory. In the Eli Whitney Metrology Laboratory, we use the combination method. Some of the personnel brought into the lab start out on the simplest type of work and advance through training to their maximum capability, and some are hired for the simpler or more routine work. This system works well in our setup because the heaviest work load is the routine type of job and the rest of the work varies considerably in amount and complexity.

To rely on aptitude testing for selection of proper personnel is somewhat hazardous. Most testing is for basic skills such as mechanical comprehension, spatial visualization, and speed and accuracy of thinking, with personality and interests inventoried. In my opinion, certain personality traits which are the hardest to clinically determine are important prerequisites for standards lab personnel. In the simpler technician areas, a sense of responsibility without over-developed aggressiveness is helpful. Generally speaking, a knowledge of a person's sense of responsibility, imagination, critical judgement and, for want of a better word, laboratory integrity would be an excellent scale to assist in selection of proper personnel. I can make no recommendation of specific personnel testing but I am sure careful evaluation of such testing can be a big assist.

The educational background of a standards laboratory technician is another area for discussion. It is easy to say the more applicable formal training, the more capable the person. However, I would like to make one observation. Where I have found an educated technician lacking in expected capability, I have found a weakness in geometry and logic. It is my personal opinion that these two subjects, per se, should be taught in grammar school. The only other observation in regards to education that I would like to make is that it is possible to hire a person with too much or too little education unless due regard is taken to the level of technology required.

One of the common systems of metrology personnel selection is that of transfer from another inspection area. Generally this works well; however, I find an occasional mismating of man and job. Frequently, an inspector wants to accept or reject and cannot understand that metrology or standards laboratory operation sometimes requires only the definition or description of a test piece. This will lead to either spending too much time in rechecking or eventual carelessness when a decision of value is required.

---

*Supervisor, Eli Whitney Metrology Laboratory, The Sheffield Corporation, Division of The Bendix Corporation, Dayton, Ohio.

Another system of selection is hiring engineers or physicists. This is a good system providing that the diversification or type of work is challenging.

One other method of employing personnel is to steal them from another company, and I don't think it would be wise to discuss this publicly.

There is a two-sided area which deserves consideration in selection of laboratory personnel, and that is motivation. It is two-sided in that it is important to know what motivates a person to want this kind of job and what will keep him motivated on the job. In the first instance, people who are typical fault-finders or perfectionists will settle in this area, and a second look at them should be taken. These people can be costly in time and money. As far as the motivation required after a technician is hired, I can only point out, as far as my own experience is concerned, that where I have been assured that I am doing an important job and am not a "necessary evil," and where I have been apprised of how my work ties in with research and development and/or the final product, I have felt the maximum motivation for continuing what is sometimes tedious or frustrating work. Also, participation in outside classes and seminars has stimulated my interest.

In conclusion, the selection of proper personnel for a standards laboratory depends on the level of technology required in conjunction with education, aptitude, personality, and motivation. All of these areas are important, and if we do not want to develop a group of mediocre laboratory personnel we should give sincere consideration to the whole technician.

There is one final comment that I must make without ulterior motive. Do not underestimate the right woman in this work at any level.

Session 6. Training of Measurement Personnel

Paper 6.3. Measurement Standards Personnel Training and Evaluation at Sandia Corporation

M. A. Elich* and H. L. Webster*

This paper outlines methods of selecting, testing, and interviewing personnel and of evaluating these methods.

Part I

Along with other responsibilities, Sandia corporation is charged with maintaining a coordinated system of physical measurement for research and development, manufacturing, and stockpile quality evaluation of nuclear ordnance. In the AEC complex, tests are required to assure that items of AEC material will function with the required degree of reliability. These tests are performed during assembly, maintenance, and surveillance of the material. A high degree of reliability is mandatory because of the ultimate purpose of the items. Reliability is achieved by maintaining and using accurate, precise, and repeatable standards, and by using standard calibration procedures.

Since quality inspection functions are performed at field test stations, it is necessary not only that the stations agree among themselves in performance, calibration procedures, and techniques, but that they also agree with previous inspecting stations, such as storage acceptance and original production. This agreement cannot be accomplished except by a rigid, coordinated system of overall instrument calibration and procedural standardization.

In order to integrate field test instrument design, calibration equipment, techniques, and procedures with overall product control, a system of standardization control was established by setting up agencies which were authorized to perform this integration.

To achieve the necessary standardization control, a method for controlling the calibration of field test instruments was established. A system primary standardization agency and such system secondary standardization agencies as needed were established to meet the needs of the field stations.

Standardization control is maintained by a technical inspection system. The system basically employs a number of traveling technical inspectors and/or equipment. These technical inspectors perform technical liaison between higher and lower echelons of standardization agencies, as, for example, between the system primary standardization agency and the system's various secondary agencies or between a secondary agency and various field stations.

The system primary standardizing agency has the following functions:

1. Maintains a system primary calibration station with suitable reference standards.
2. Maintains portable standards which are transported periodically to the system secondary standardizing agencies.
3. Provides system primary calibration services for the secondary agencies either directly or through traveling inspectors.
4. Is responsible for the coordination of procedures and techniques among the field stations insofar as they are related to field test instrument standards.
5. Is responsible for the coordination among the system secondary standardizing agencies.

The system secondary standardizing agencies have these functions:

1. Maintain system secondary calibration stations with suitable reference standards.
2. Maintain portable standards which are transported periodically in the various field stations by traveling inspectors.
3. Provide calibration services for the field stations either directly or through inspectors.

The field stations (1) maintain such field working standards as may be required and (2) provide calibration service directly or indirectly for testers.

---

*Sandia Corporation, Albuquerque, New Mex.

With the establishment of the AEC system of standardization, it became necessary to train military, AEC, and Sandia Corporation personnel in the theory, techniques, and procedures of measurement standards so that the complete standardization program could be realized.

Training has been accomplished for civilian and military supervisors, instructors, engineers, and technicians. In setting up the course of instruction for military depot level personnel, it was assumed that the training would be for new hires or personnel who had no experience in measurement standards. Later, refresher training would be offered. In this case, the training would be such as to lead to a certification of competence if requested by the Military. However, certification of trainees was never requested, so the planned course of study was shortened with no follow-up evaluation.

The content of the depot level course as outlined in the enclosure covered cross-checks of such reference standards as standard cells, potentiometers, volt boxes, shunts, resistors, mercury barometers, and the calibration of several transfer standards.

Approximately twenty-five depot level courses were offered to Army, Navy, and Air Force personnel from 1953 to 1959; at this time, the AEC field tester calibration responsibility was assumed by all DOD agencies.

The depot level course was geared to military technician level. Since the trainee's experience and qualifications were not known until the course convened, the method of presentation and the material covered were not necessarily the same from class to class. The conduct of the class was usually as follows.

The standardization system was described from the standpoint of electrical and pressure standards. The flow of information was traced from the National Bureau of Standards to the field tester. The theory of electrical and pressure standards instruments, their capabilities, and their limitations were presented in a lecture-type presentation. The laboratory phase of the instruction consisted of the actual crosscheck and calibration of standards equipment by the student, using "cook book" procedures. The procedures used were published by Sandia Corporation in the AEC TP T-Series which was directive on the military. Since the DOD agencies have integrated the above described standardization system into their own system, the calibration procedures are no longer directive on the military by the AEC. The classes lasted for two weeks and were always small in size (four to six students). The class was divided into two groups. While one group received instruction in pressure standards, the other group was instructed in electrical standards. This arrangement made the instructor-student ratio very good and provided almost continuous individual instruction. An examination was given at the conclusion of the course. The test was designed primarily as a teaching tool rather than as an evaluation and merit-measuring instrument. A certificate of course completion rather than competence was awarded.

In the case of the field level course, which lasted one week, the instruction was much the same as the depot level course as far as student participation was concerned; however, the cross-check of reference standards and calibration of transfer standards were demonstrated. Most of the time was spent by the students in actually calibrating different field testers under close instructor supervision.

With the establishment of secondary reference laboratories at Sandia Corporation Field Inspection Area Offices, there was a need for more calibration training at the secondary level. A course was set up for four field office electrical inspectors. This course was much more extensive than that given to the military depot level personnel. The course lasted five weeks. The amount of theory presented and the types of instruments calibrated were much more extensive, as can be noted from the course outline. In the short time scales, it was not possible to train personnel adequately in all phases of secondary laboratory requirements. Because of this, training effort was concentrated on the basic techniques of calibration, especially in the area of interconnecting and correctly using sensitive instruments. It was felt that a good basic knowledge of sensitive instrument theory, capabilities, limitations, and operating techniques would carry over to other types of instruments. To aid the inspectors, who were beginners in the field of measurements standards even after a basic course of instruction, it was felt desirable and necessary to provide them with step-by-step calibration procedures for cross-checking their reference standards and for calibrating numerous transfer standards. To provide the formal cross-check and calibration procedures for secondary reference and transfer instruments, a manual, *Calibration of Secondary Standards for Product Test Equipment* (SM 6-3), was published. This document presently covers the cross-checking of such standards as standard cells, potentiometers, volt boxes, shunts, resistors, and the calibration of such transfer standards as D'Arsonval-type meters, differential volt meters, and electrostatic volt meters. The manual will be revised and supplemented as required.

Experience from the various calibration courses has pointed out the difficulty of knowing or anticipating individual training needs. Little foreknowledge of the specific qualification, experience, and background of each student was available. This forced the instructor to "play by ear" rather than simply present a rigidly organized course. A rigorous liaison program could eliminate this difficulty if depot level courses are again requested by the military or any AEC agency.

As was pointed out before, no trainee has been certified because no request for certification was received from the military. Personnel can not be certified as calibrators unless:

1. education qualifications and experience are controlled,
2. length and content of course are controlled,
3. the trainer conducts practical examinations and audits, and has a degree of command control for follow-up evaluation, and
4. the certification is specific with respect to the area in which the individual is competent.

Since the trainees were not certified, no formal follow-up evaluation of the personnel who completed the course was made. However, Sandia Corporation Field Inspection Area Office personnel have been evaluated by means of audits conducted by the system primary laboratory. This type of evaluation will continue.

The program for evaluating measurement standards personnel at Sandia Corporation has not been as formal as the program for training personnel. However, there is a continuing need to answer questions like the following:

1. What is expected of me?
2. Why am I not told how well I am doing?
3. How do I get ahead?
4. What should I be learning now to prepare myself for advancement?
5. Does management know what I can do, what I am doing, or why I am doing it?

The answers to these questions should be found in a sound performance evaluation program. The primary objectives of such a program for professional and semiprofessional standards personnel are generally considered to be:

1. Determine present objectives and status of work progress,
2. Provide realistic and equitable measurements of value of services for wage and salary purposes,
3. Aid in selecting individuals for promotion, transfer, or separation, and
4. Assist in motivating employee toward improvement.

Evaluation is difficult because work effectiveness does not yet lend itself to precise measurements. An employee's performance cannot be fairly evaluated until both he and his employer are in agreement as to what the job requirements are. The first step is to define the job. It must be kept in mind that it is not the individual who is being evaluated; it is his performance as related to his work responsibilities. Personality traits such as drive, work potential, compatibility, loyalty, creativity, self-reliance, enthusiasm, persistance, responsibility, efficiency, reasoning ability, and emotional maturity describe the individual; they do not necessarily show how effectively he discharges his responsibilities. If we go back a few years to the classroom, we may recall that such traits had little direct affect in obtaining a satisfactory rating on a final examination. Performance in a calibration laboratory is quite obviously more difficult to measure.

Work objectives probably should be determined independently by supervisor and employee. The argument for asking the employee to write down his job objectives are as follows:

1. The supervisor is given insight into the level at which the employee thinks.
2. The supervisor is given a clearer understanding of the job.
3. The employee can see himself with more perspective than he might otherwise.
4. The employee is aided in remembering objectives until the next review because he has taken the time to spell them out himself.

It must be assumed that finally the supervisor and employee agree on objectives. The specific problem of the evaluating of performance in a standards laboratory might include the following:

1. Keeping a limited running record of job objectives (and adequacy of results) for each employee.
2. Keeping a limited running record of measurement audit results. One procedure tried in our environmental standards section is to have standards recently calibrated by NBS calibrated by our laboratory personnel as for a customer, including a final report or certificate. At this point, and not before, the NBS values are revealed and compared. A few notes are usually adequate for review at the next employee evaluation.

The evaluation must be accurately related to the employee's pay. Even if money, in itself, is not important to the employee, it will serve as an indicator to him as to how well he is performing. A salary reduction for an employee following a very favorable performance evaluation interview would undoubtedly have a negative effect upon his motivation toward improvement.

The employee's performance has value. The value of raw material in the ground is usually a very small portion of the cost of a final product. An auto or a refrigerator is primarily an exhibit of packaged labor. In a free economy, prices are controlled in the long run by supply and demand. Generally, in the U. S. labor market for professional and semiprofessional technical personnel, communication is so good that no great strain is put upon the technical supervisor in the establishment of the pay range for a given service. If the performance evaluation by the supervisor is accurate, the employee's salary can be established at a level compatible with the labor market inside and outside of the company or organization involved.

It must be concluded that unless the supervisor and employee both have an accurate up-to-date understanding of the job objectives, evaluation of performance becomes extremely difficult and probably quite unfair.

Secondary Reference Calibration Course Outline

I. Introduction
    A. Requirements for calibration within special weapons organization
    B. Training responsibility of Sandia Corporation
    C. Course of study
    D. Class organization
II. Pressure, Temperature, and Humidity Calibration
    A. Pressure calibration
        1. Baroswitch assembly function
            a. Baro display and explanation of component function
            b. Explanation of baro application to Special Weapon Systems
        2. Display of all pressure calibration and pressure test equipment
            a. Primary reference
            b. Primary transfer
            c. Secondary reference
            d. Secondary transfer
            e. Local reference
            f. Component test
        3. Explanation of the method by which National Bureau of Standards information is carried down through the Special Weapons Calibration system to field tester and finally to the weapon. Ap-

proximate accuracy at each calibration level including weapon components.

4. Manuals
   a. Calibration procedure manuals
   b. Reporting system manuals (TP-5-11)
5. U-tube differential mercury manometer
   a. Definition of barometer, manometer, differential mercury manometer
   b. Definition of specific weight, density
   c. Pressure variation in a confined fluid at rest
   d. Variation of specific weight of a material with its temperature
   e. Variation of specific weight of a material with gravity to which exposed
      (i) Systems of units (force, mass, and acceleration)
         (a) English gravitational and absolute
         (b) Metric gravitational and absolute
   f. Variation of gravity with evaluation, latitude, and local earth composition (strata)
6. "Barometer"
   a. Reading the "barometer"
   b. Need for temperature and gravity corrections
   c. Operation of temperature and gravity correction mechanism
   d. Calibration procedure
7. Barometer to secondary transfer standard
   a. Display of altimeter in disassembled condition and explanation of its operation
      (1) Differences between types of altimeters
      (2) How altimeter is read
   b. Barometer to pneumatic schematic
      (1) Explanation of vibrator, calibrate and bleed tanks and calibrate and bleed valves
   c. Barometer to secondary transfer standard calibration procedure
8. Secondary transfer standard to tester (one check point)
9. Tester to baroswitch assembly
10. Barometer to baroswitch assembly
11. Calibration of pressure gages
12. Written and practical test

III. EMF Calibration
   A. Secondary reference calibration
      1. General discussion of potentiometers and accessories including elementary series and parallel circuits, Section A of TP-T-51 reporting system, and operation and care of equipment, e.g., Power Supplies, Galvos, and Counter.
      2. Actual practice in use of potentiometer by performing cross-checks in TP-T-51 (Sections B through P)
         a. Practice in setting up and reading
         b. Cross-check of standard cells
         c. Cross-check of standard resistors of same nominal values and of different nominal values. The latter requires the application of POT corrections.
         d. Cross-check of potentiometer (includes volts dial, slide wire, and multiplier)
         e. Check of volt box and multirange shunt
   B. Secondary transfer equipment calibration
      1. General discussion of operation and care of Polyranger (including D'Arsonval principle, thermocouple, accuracy and precision meaning, reading scales, calibration card, and use of correction factors)
      2. Actual calibration of Polyranger with potentiometer and accessories on applicable scales (both ac and dc voltage and current)
      3. General discussion of operation and care of Wheatstone Bridges test set (including use in Murray Loop position, normal resistance measurement position, plus care and operation of galvos)
      4. Actual Wheatstone Bridge calibration
      5. General discussion of portable frequency meters, standard megohm resistors, stop watches, and T-500 equipment
      6. Stop watch calibration with counter

IV. Field Trip Through Sandia Primary Laboratory

V. New Developments in T-500 Series Equipment

Course Outline for Secondary Reference and Secondary Transfer Standards

I. Introduction
   A. Requirement for Standardization
      1. History
      2. Philosophy
      3. Scope
   B. Calibration
      1. Equipment
      2. Procedures
      3. Techniques
   C. Course of Study
      1. Agenda
II. Definitions
   A. Standardization
   B. Calibration
   C. Calibration Loop
   D. Validity
   E. Accuracy
   F. Precision
   G. Comparison Operations
      1. Direct Method
      2. Transfer Method
   H. Invalidation
III. The Art of Measurement
   A. The Choice of Method
   B. The Choice of Equipment
   C. Accuracy and Precision
   D. Errors in Measurement
   E. Laboratory Practices
      1. Procedures
      2. Techniques
IV. Electrical Standards
   A. Secondary Reference
      1. Potentiometer and Standard Cell Methods
         a. The Constant-current Potentiometer
            1. Operating Principles

      - 2. Simplified Circuits--Theory
    - b. Types of Potentiometers
      - 1. Simplified Circuits--Theory
      - 2. Operation
      - 3. Capabilities--Limitations
      - 4. Accuracy
    - c. Standard cell (Voltage Standard)
      - 1. Standard--Unsaturated
      - 2. Use, Limitations, Accuracy
    - d. Null Indicators (Galvanometers)
      - 1. Shadow Boxes
      - 2. Microvolt Amplifiers
      - 3. Other Types of Indicators
    - e. Shunts
      - 1. Galvanometer
      - 2. Ayrton
      - 3. Heavy Current
      - 4. Multirange
    - f. Voltage Dividers
      - 1. Volt Boxes
        - a. Multiplying Factor
        - b. Use--Limitations--Accuracy
  - 2. Cross-Check of Standard Cells With Potentiometer
    - a. Setting up Potentiometer
    - b. Working Battery
    - c. Dial Reading Practice
    - d. Procedures
    - e. Techniques
  - 3. Cross-Check of Standard Resistors
    - a. Techniques and Procedures
    - b. Of Same Nominal Value
    - c. Of Different Nominal Value
  - 4. Cross-Check of Potentiometer
    - a. Techniques and Procedures
    - b. Volts Dial
    - c. Slide Wire
    - d. Multiplier
- B. Secondary Transfer
  - 1. Power Supplies
    - a. Theory of Operation
    - b. Application
    - c. Capabilities and Limitations
    - d. Accuracy
  - 2. Electrical Indicating Instruments (General)
    - a. DC Permanent Magnet--Moving Coil
    - b. DC Voltmeters
    - c. DC Ammeters
    - d. Electrodynamometer Voltmeters Ammeters
    - e. Moving Iron Instruments
    - f. Rectifier Instruments
    - g. Thermocouple Instruments
  - 3. Polyrangers
    - a. Capabilities--Accuracy--Use
    - b. Calibration with Potentiometer
      - 1. Procedures and Techniques
  - 4. Resistance Measurements
    - a. Standard Resistors
    - b. Types of Resistors
    - c. Resistance Measurements
      - 1. Ammeter - Voltmeter Method
      - 2. Ohmmeters
      - 3. Comparison Method
    - a. Potentiometer Method
    - b. Wheatstone Bridge
      - 1. Theory
      - 2. Sensitivity
      - 3. Application
      - 4. Limitations--Capabilities--Accuracy
      - 5. Checking by Potentiometer
    - c. Kelvin Bridge
      - 1. Theory
      - 2. Sensitivity
      - 3. Application
      - 4. Limitations--Capabilities--Accuracy
    - d. High Resistance (Megohm) Measurements
  - 5. Electrostatic Voltmeters
    - a. Theory of Operation
    - b. Types
    - c. Applications
    - d. Capabilities--Limitations
    - e. Sensitivity--Accuracy
    - f. Operating Procedures
  - 6. Differential Voltmeters
    - a. Theory of Operation
    - b. Types
    - c. Applications
    - d. Capabilities--Limitations
    - e. Sensitivity--Accuracy
    - f. Operating Procedures
  - 7. Portable Potentiometers
    - a. Theory of Operation
    - b. Types
    - c. Applications
    - d. Capabilities--Limitations
    - e. Sensitivity--Accuracy
    - f. Operating Procedure
  - 8. Capacitance and Inductive Bridges
    - a. Theory of Operation
    - b. Types
    - c. Applications
    - d. Capabilities--Limitations
    - e. Sensitivity--Accuracy
    - f. Operating Procedures
  - 9. Calibrators
    - a. Theory of Operation
    - b. Applications
    - c. Capabilities--Limitations
    - d. Sensitivity--Accuracy
    - e. Operating Procedures
  - 10. Oscilloscopes
    - a. Theory
    - b. Simplified Circuits
    - c. Capabilities--Limitations
    - d. Applications
    - e. Operating Procedures and Techniques
  - 11. Counters
    - a. Theory
    - b. Simplified Circuits
    - c. Capabilities--Limitations
    - d. Applications
    - e. Operating Procedures--Techniques
  - 12. Vacuum Tube Voltmeters
    - a. Theory of Operation
    - b. Types
    - c. Applications
    - d. Capabilities--Limitations
    - e. Sensitivity--Accuracy
    - f. Operating Procedures
  - 13. Microampere Measurements
    - a. Using Potentiometer
    - b. Procedures and Techniques
  - 14. Pulse Generators
    - a. Theory of Operation

        - b. Types
        - c. Applications
        - d. Capabilities--Limitations
        - e. Sensitivity--Accuracy
        - f. Operating Procedures

- V. Pressure Standards
    - A. The "U" Tube Mercury Manometer
        - 1. Principles
        - 2. Definitions
            - a. Standard Pressure
            - b. Standard Temperature
            - c. Standard Gravity
        - 3. Effects of Temperature and Gravity
        - 4. Pressure Altitude
            - 1. NACA Report 538
        - 5. Absolute, Gage, Barometic, Vacuum Pressure
    - B. The Mercury Barometer
        - 1. Principles
        - 2. Capabilities--Limitations
        - 3. Applications
        - 4. Accuracy
        - 5. Temperature and Gravity Corrections
        - 6. Reading the Barometer
    - C. Hass to Hass Calibration
        - 1. Procedures
        - 2. Techniques
    - D. Altimeters
        - 1. Principles
        - 2. Calibration Using Barometers
            - a. Procedures and Techniques
    - E. Pressure Gages
        - 1. Calibration Using Barometer
            - a. Procedures and Techniques
    - F. Dead Weight Tester
        - 1. Principles
        - 2. Calibration

Session 6. Training of Measurement Personnel

Paper 6.4. Selection, Training, and Evaluation of Precision Measuring Personnel in the Air Force

J. N. Eversole*

1. Introduction

Efforts to provide a single integrated calibration program throughout the Air Force were begun in the late 1950's. With the advent of new and more complex armament and weapons systems, it became evident that a career area to support calibration and certification of precision measuring equipment would have to be developed. In 1958 a committee comprised of USAF representatives and all major air commands was formed to devise a single integrated program. The intent of the original program was to support armament systems, but the concept changed rapidly to support all test equipment in all weapons systems, armament systems, and missile areas, along with other functions which require the use of precision measuring equipment.

Because of the transportation time involved and the difficulties of shipping precise instruments to and from the depots, NBS, and other reliable calibration facilities, it was decided that 160 Air Force bases should be given the capability to provide calibration service to the precision measuring equipment (PME) within their prescribed geographical area. The most important consideration in the program was to place the capability for calibration and certification of accuracies as close to the weapon or system as possible. Therefore, each Air Force base was given the mission of building facilities and gaining a capability for support of the precision measuring program. Because of the wide dispersal and location of bases from the Air Materiel Areas (AMAs), depots, NBS, and other dependable calibration support, self-sufficiency of the base laboratories is of paramount importance. Although the bases still depend upon the AMA's and depots to certify the accuracy of their standards periodically, the responsibility for accuracies of the weapon and/or system has largely been relegated to the Air Force base laboratories. The accuracies of the entire calibration system are directly traceable to NBS.

2. Personnel Selection

Personnel to support this program was the first and primary consideration. For this new program no personnel resources were available in the Air Force who had experience in calibration of precision measuring equipment. Therefore, it was necessary to develop a complete career program from scratch.

The selection of personnel emphasized the importance of electronics training and experience, since 90 percent of the calibration tasks were of the electronic variety. Consideration was given to the selection of personnel who were already semiskilled or skilled in a lateral technician area. It is important that the personnel entering this field have extensive electronics backgrounds, have high aptitude quotients in both electronics and mechanics, and have experience in maintenance of test equipment.

To assure that the commands were providing the training course with qualified personnel, a preentry examination was made mandatory for selection of personnel to attend training. The bases administer the test under strict control to assure that the man is qualified before sending him to the training. A 150-question test comprised largely of mathematics, electronics, and physics is used for this preentry selection. A cutoff of 90 correct out of 150 items is required for qualification to enter training. The test has been a reliable predictor of success of students entering the course and attrition from training has been held to allowable limits.

These are the prerequisites established for the personnel entering the training conducted by the Air Force at Lowry AFB, Colo.

---

*Lowry Air Force Base, Colo.

3. Training

In April 1959 a course was begun at Lowry AFB to train personnel in the use and maintenance of Air Force calibration standards. The first course placed primary emphasis on maintenance of test equipment. This concept was changed quickly to one of use of the calibration standards and calibration of the standards themselves. Step-by-step (cookbook) procedures using 33K series (USAF calibration procedures) technical orders were taught throughout the course. The course was designed to consist of approximately 80 percent electronic and 20 percent physical standards. The allocation of time and emphasis was due to the fact that the majority of USAF calibration jobs consist of electronic measurements. This concept, too, has changed to some degree and more emphasis is now being placed on physical measurements.

Emphasis in Air Force training is generally placed upon performance--using the actual equipment. Due to the continuous growth of the precision measuring equipment field to encompass new equipment and the demand for higher accuracies, the training course was plagued with continuous changes. In January 1962 the course at Lowry AFB was revised to place major emphasis upon the principles of measurement.

To avoid the probem of changing standards, the course would be PRINCIPLE oriented, since only the standards change--not the principles.

The training would endeavor to make professional metrologists of the students as a goal to work toward.

It would emphasize basic principles and fundamental understanding, without a lessening in the amount of practical experience on the items themselves.

It would include all new standards which have been added to the list, plus the principles of anticipated future standards. Here, trainers would be used in place of the physical standards.

The standard would be taught, not as an *item*, but as a trainer to demonstrate the principle of measurement associated with that standard.

The primary purpose of our training is to supply the laboratories with the most dependable personnel available in the Air Force. These personnel, with a minimum of supervision, are expected to make accurate measurements and be relied upon for certification of the accuracies demanded.

Due to the selection of senior airmen, the career area of precision measuring equipment is attractive to most of the personnel in it. It has been given and has earned prestige throughout the Air Force. The complexity of the career area has been recognized by providing the personnel in it with proficiency pay, which is the addition to the normal pay and allowances.

4. In-Training Evaluation

During the training, aside from personal evaluation by the instructors and supervisors, after each homogeneous grouping of material or equipment, tests are given to the students. These tests are both written and performance-type examinations. The performance-type examination is generally a calibration procedure wherein the student is required to perform calibrations of test equipment in accordance with the applicable T.O. to tolerances specified by the training school. The equipment tolerances specified by the technical orders are adapted to the classroom situation.

5. Evaluation of Laboratories and Personnel

Methods of evaluating the performance of personnel in base laboratories have been talked about at length. The system proposed for use by the USAF Calibration Division is contained in T.O. 33-1-14. This T.O., in addition to outlining the USAF calibration program, establishes a method for certification of laboratories. An extract from T.O. 33-1-14 follows. Quote: "The USAF Calibration Division will establish a system for monitoring all AMA calibration laboratories and to certify their ability to perform calibration on each of the measurement categories. This monitoring system will consist of the following:

a. Prepare test problems in specific measurement categories and mail or deliver the same type problem to each AMA Laboratory.

b. AMA Laboratory personnel will complete the problem and provide the complete data on the measurement test including procedures, equipment used, and environmental data.

c. Solutions will be analyzed by the USAF Calibration Division and the results published to each AMA. AMAs producing results within the established tolerances will be certified in the particular measurement category.

d. When results do not meet established tolerances, the AMA will be notified immediately that they did not qualify. The USAF Calibration Division will provide assistance to the AMA to identify and rectify the cause so the AMA Laboratory can be certified as quickly as possible.

The USAF Calibration Division will establish a program for each AMA supporting base PME Laboratories to "problem test" the Base PME Laboratories and certify those successfully completing test problems." Due to lack of personnel and funds, the evaluation of base laboratories throughout the Air Force has been limited. Two methods are used by the Air Training Command for evaluating performance of the graduates. The first is a use of graduate evaluation forms which are sent to the supervisor of the graduates asking for evaluation of their performance. The second is visits by technical personnel and instructors to base laboratories for a firsthand look at the performance of the graduates.

Due to lack of personnel and funds, the evaluation of base laboratories throughout the Air Force has been limited. Two methods are used by the Air Training Command for evaluating performance of the graduates. The first is a use of graduate evaluation forms which are sent to the supervisor of the graduates asking for evaluation of their performance. The second is visits by technical personnel and instructors to base laboratories for a firsthand look at the performance of the graduates.

Another method used throughout the Air Force for self-evaluation of the efficiencies of laboratories is the AFM 66-1 Maintenance Data, Collection Information. This system can be used to determine downtime and turnaround time of equipment within laboratories. Through a system of maintenance manhour accounting, it is possible to determine the proficiency of the laboratory.

Because of the great numbers of personnel and laboratories involved, no single system of evaluation is in effect. The most reliable--but certainly least desirable--is the "after the fact" evaluation where the proficiency is judged by the percentages of mission success and failure. The Air Force is continuously working on evaluation procedures which will reliably assure the success of the mission. Much of this reliability is contingent upon the selection, training and evaluation of personnel performing precision measuring jobs in support of the mission.

1

Session 6. Training of Measurement Personnel

Paper 6.5. A Precision Electrical Measurements Course

R. E. Travis*

This paper describes how a precision electrical measurements course was developed from the information contained in courses taught by the University of California at Los Angeles, the George Washington University at Washington, D. C., and other sources. This course consists of 34 two-hour sessions.

1. Introduction

At the time of submission of this paper, the minutes of the National Conference of Standards Laboratories Ad Hoc Committee Meeting held in Los Angeles last September had not yet been published. It was surprising to note recently in Mr. Leon Hachey's summary of the questionnaire[1] published in these minutes that the major field of interest was management and administration of AC (low frequency) and DC electrical measurements (102 out of 169 replies). However, in the next question, "problems which the proposed association of standards laboratories should help solve," personnel training was listed on only 6 out of 169 replies.

This summary indicates one of the following conditions: 1. Low frequency and DC laboratories have no training problem. This would imply also that they have no personnel problem. 2. The training problem is little recognized by management represented at the September 1961 conference. 3. The conference attendees put little faith in the Association of Standards Laboratories' being able to provide a solution to training problems.

Our experience at the Boeing Primary Standards Laboratory does not verify the first condition derived from the summary. If we now proceed on the assumption that condition number one is not the true position of many of our laboratories, we can profitably seek solutions to the other conditions that were derived from the survey.

Condition number two is a common management problem that exists when management is not technically involved in the laboratory function. Here the solution is in technically directed management training. This is one area where the NCSL can be of great benefit. This conference of standards laboratories here this week, and the International Conference on Precision Electromagnetic Measurements next week, help to provide technical training for the standards laboratory management people who participate. Easily attended regional seminars can and should be organized to accelerate our learning process. Laboratory management should actively participate in the local training program in both the organization and presentation. To aid in this participation, the specialized short courses such as we will mention should be attended.

The third condition that was indicated by the summary can be improved by an active sharing of information that is already in existence in the laboratories represented here today. It is with this in mind, and in line with Mr. William Wildhack's[2] request that the NCSL be an organization wherein our standards laboratories pool their experience, that we decided to present this information here.

The Boeing Precision Electrical Measurements Course was based upon training received at courses taught by NBS personnel and supplemented by some excellent references (appendix I) in the electrical measurements field. The Low Frequency

---

*The Boeing Company, Seattle, Washington.

[1] Minutes of the Laboratory Conference of the Ad Hoc Committee on an Association of Standards Laboratories held in Los Angeles, September 15, 1961 (published June 4, 1962).

[2] Letter 30.60 from William A. Wildhack, Associate Director of NBS, to Frank McGinnis, AIA Quality Control Committee, dated Nov. 3, 1961.

Electrical Group of the Boeing Primary Standards Unit was able to send one engineer to the UCLA course[3] taught by Drs. Thomas and Harris in 1960, and another to the same course taught by Dr. Harris and Peterson in 1961. Another engineer attended the GWU course taught by Dr. Harris, Peterson, and McNish last summer.

Desiring to disseminate this information to all the personnel in the group, several of whom were new and inexperienced in precision measurements, we established a formal training course. The Boeing Industrial Training Unit was invited to participate. They provided one instructor who participated as a student. He will assist in future training requirements of a similar nature and will extract pertinent information for inclusion in other courses of their electronic training program. They also provided the formalized class status, including the training charge, so the course could be held as paid-time training.

Two hours per week was considered the maximum that could be allotted to the training. Sessions were scheduled to include our personnel from both work shifts. The available material was covered in one year, not including the summer months. Instruction was divided among the engineerin staff. Since we were understaffed at this time special emphasis was given to the subject assignments so as to require a minimum of preparation time. To prevent undue hardship on the engineers or excessive interference with their normal work, the schedule attempted to allow intervals of three to four weeks between presentations by any individual. This caused some sacrifice of the most desirable course sequence. (See appendix II. An improved sequence is shown in appendix III.)

Session preparation included a complete draft of the subject to be covered. This was considered the best way to consolidate our previous classroom notes, several texts, and the experience of the individual engineers. This gives us a complete course with text, all or part of which can be presented at a later date to new people in our own laboratory, or to the personnel of other laboratories or divisions within the company. One instructor can now present the course with a minimum of preparation time.

The individual sessions follow the general outline shown in Chart 1, modified for the material available. Occasionally two sessions are used to cover the material of one outline.

TITLE

I. INTRODUCTION AND HISTORY
II. THEORY
III. DESIGN AND CONSTRUCTION
IV. CALIBRATION AND USE
V. REFERENCES

Chart 1.

One of our recommendations for future sessions (or to others who may use the course) is to limit the size of the class to ten (10). Preparation for future presentations would emphasize more training aids and classroom problems. Homework problems were little used in this first go-around, but this is also an appropriate part of such a course.

This presentation here today may to some extent be considered a tribute to the dedicated scientists from NBS and the competent men from industry who have made possible the highly specialized short courses mentioned at the beginning of this per. It is hoped that these courses will continue be offered and that other courses will be added broaden the field of study within the standards oratories requirements. Even with the full ticipation of industry in sending selected individuals to these courses, it must be realized that the costs involved and facilities available prevent these activities from being a solution to the training problem for our personnel. Only as we digest and disseminate the information obtained from a concentrated short course can we obtain the full benefits from an individual's participation in that course.

It is hoped that our experience presented here will help other standards laboratories in the accomplishment of their training requirements. For those specifically interested in the low frequency electrical measurements field, we have a limited number of copies of the complete course as outlined in the schedule (appendix II).

For copies of this course contact Charles Johnson, Primary Standards Metrology Laboratory Chief, The Boeing Co., P. O. Box 3707, Seattle 24, Washington.

Appendix I. Book Reference List

1. Hague, B., Alternating Current Bridge Methods, fifth edition (Pitman, 1959).
2. Harris, F. K., Electrical Measurements (Wiley & Sons, Inc., New York, N.Y., 1952).
3. Laws, F. A., Electrical Measurements, second edition (McGraw-Hill Book Co., 1938).
4. LePage, W. R., Analysis of Alternating Current Circuits (McGraw-Hill Book Co., 1952).
5. National Bureau of Standards Handbook 77, Volume 1 (U. S. Government Printing Office, 1961).
6. Partridge, G. R., Principles of Electronic Instruments (Prentice-Hall, Inc. 1958).
7. Stout, M. B., Basic Electrical Measurements, second edition (Prentice-Hall, Inc., Englewood Cliffs, N.J., 1960).

---

[3] Reliability and Statistical Methods in Industry--Basic Electrical Measurements Course; University of California, Los Angeles; E. P. Coleman, coordinator.

[4] The Summer Program in Measurement Science, Basic Electrical Measurements; George Washington University, Washington, D.C.; 1961.

Appendix II.

To: Primary Standards Low Frequency Electrical Group

cc: C. E. Johnson
H. Heighton
M. Strum
L. Bowen
R. Weir

Subject: Precision Electrical Measurements Course, Revised Schedule

Reference: 2-4823-1-176, Course Schedule dated Sept. 25, 1961

The Low Frequency Electrical Group of the Primary Standards Unit is conducting special training in Precision Electrical Measurements as scheduled below. Classes will meet from 1:30 to 3:30 P.M. in conference room 12C23 (or other area as posted on the 12C23 door) on Wednesdays shown.

Session Number	Instructor	Topic	Date
			1961
1	Seefeld	Philosophy of Measurements	Sep. 27
2	Barber	Calculations (Algebra of Small Quantities, Proportional Parts, Parts per Million, Percentage)	Oct. 4
3	Barber	Resistance Measurement, Wheatstone Bridge, DC detectors.	Oct. 11
* 4	Travis	Galvanometers, galvo damping and mounts	Oct. 18
5	Zugel	AC Bridges, Sources, Detectors, Shielding, "j" algebra, balance equations	Oct. 25
* 6	Seefeld	Standard Resistors, Construction, Stability, Temperature Coefficient; Series and Parallel Combinations.	Nov. 1
** 7	Barber	Standard Resistor Comparator and Oil Bath. Calibration of Comparator and Standard Resistors	Nov. 8
* 8	Zugel	The Direct Reading Ratio Set, Waidner-Wolff Elements, Exact Ratio	Nov. 15
** 9	Travis	Calibration of the DRRS	Nov. 29
**10	Seefeld	Calibration of Wheatstone Bridge with DRRS	Dec. 6
* 11	Barber	Kelvin Double Bridge Operation and Calibration	Dec. 13
* 12	Zugel	Phase Measurement, Part I; Gertsch & Boeing Systems	Dec. 20
* 13	Seefeld	Resistance Measurement Summary and Very High Resistance Measurement	Jan. 3 1962
* 14	Travis	Standard Cells and Zener Diode Voltage References	Jan. 10
* 15	Seefeld	Potentiometers, Measurement of Small DC Voltages, Lindeck Element. Standard Cell Comparator	Jan. 17
* 16	Barber	Universal Ratio Set Calibration and Use	Jan. 24
* 17	Zugel	AC Voltage and Current Transfer Devices	Jan. 31
* 18	Seefeld	Measurement of DC Voltage, Voltage Multipliers, AC and DC High Voltage Facility	Feb. 7
* 19	Travis	Volt Box Calibration and Use. NBS Type Volt Box	Feb. 14
**20	Zugel	Boeing Standard Divider and Calibration of Kelvin-Varley Dividers	Feb. 21
* 21	Barber	Measurement of DC Current, Current Shunts, High Current Supply and Load Bank	Feb. 28

Session Number	Instructor	Topic	Date
22	Driscoll-Steinmetz	Measurement of AC Voltage, Current and Power; AC Indicating Instruments	Mar. 7
* 23	Seefeld	AC & DC Indicating Instrument Repair and Calibration; Magnet Adjustment	Mar. 14
**24	Barber	Resistance Thermometer and Thermometer Bridge Calibration	Mar. 21
* 25	Zugel-Park	Instrument Transformers, Current and Voltage. Current Transformer Test Set	Mar. 28
* 26	Travis	Ratio Transformers, Theory, Design and An Application	Apr. 4
* 27	Seefeld	Standards of Inductance and Capacitance	May 16
28	Park	AC Resistance Standards and Measurement	Apr. 18
**29	Zugel	Capacitance Calibration, Ratio Arm Bridge, AC Bridge Grounds	Apr. 25
30	Seefeld	Hay Bridge, Owen Bridge and Schering Bridge	May 2
* 31	Barber	Boeing Limit Bridge and Wenner-Barber Ratio Set	May 9
32	Driscoll	Power Factor Measurement	Jun. 6
**33	Zugel	Mutual Inductance Measurement, Maxwell-Wien Bridge	May 23
34	Park	Phase Angle Measurement, Part II	May 29

R. E. Travis,
Supervising Engineer
Primary Standards Unit
Low Frequency Electrical Group

* Demonstration
** Experiment

Appendix III.

Suggested Course Sequence For Precision Electrical Measurements Training Course

Session Number	Topic
1	Philosophy of Measurements
2	Calculations (Algebra of Small Quantities, Proportional Parts, Parts per Million, Percentage)
3	Resistance Measurement, Wheatstone Bridge, DC Detectors
4	Galvanometers, galvo damping and mounts
5	Standard Resistors, Construction, Stability, Temperature Coefficient; series and parallel combinations
6	Standard Resistor Comparator and Oil Bath. Calibration of Comparator and Standard Resistors
7	The Direct Reading Ratio Set, Waidner-Wolff Elements, Exact Ratio
8	Calibration of the DRRS
9	Calibration of Wheatstone Bridge with DRRS
10	Kelvin Double Bridge Operation and Calibration
11	Universal Ratio Set Calibration and Use
12	Resistance Thermometer and Thermometer Bridge Calibration
13	Resistance Measurement Summary and Very High Resistance Measurement
14	Boeing Standard Divider and Calibration of Kelvin-Varley Dividers

Session Number	Topic
15	Boeing Limit Bridge and Wenner-Barber Ratio Set
16	Standard Cells and Zener Diode Voltage References
17	Potentiometers, Measurement of Small DC Voltages, Lindeck Element. Standard Cell Comparator
18	Measurement of DC Voltage, Voltage Multipliers, AC and DC High Voltage Facility
19	Volt Box Calibration and Use. NBS Type Volt Box
20	Measurement of DC Current, Current Shunts, High Current Supply and Load Bank
21	AC Voltage and Current Transfer Devices
22	Measurement of AC Voltage, Current and Power; AC Indicating Instruments
23	AC & DC Indicating Instrument Repair and Calibration; Magnet Adjustment
24	AC Bridges, Sources, Detectors, Shielding, "j" Algebra, Balance Equations
25	Instrument Transformers, Current and Voltage. Current Transformer Test Set
26	Ratio Transformers, Theory, Design and An Application
27	Standards of Inductance and Capacitance
28	AC Resistance Standards and Measurement
29	Capacitance Calibration, Ratio Arm Bridge, AC Bridge Grounds
30	Hay Bridge, Owen Bridge and Schering Bridge
31	Mutual Inductance Measurement, Maxwell-Wien Bridge
32	Phase Measurement, Part I; Gertsch and Boeing Systems
33	Phase Angle Measurement, Part II
34	Power Factor Measurement

NBS Miscellaneous Publication 248
Proceedings of the 1962 Standards Laboratory Conference

Session 6. Training of Measurement Personnel

Paper 6.6. The Navy Calibration Training Program

Stanley Crandon*

This paper describes the Navy's calibration training program, which covers the areas of electrical measurements, electro-mechanical measurements, and linear and optics measurements, including the basic fundamentals and principles used to provide familiarity with standards and associated equipment.

1. Introduction

In the early stages of the Navy Calibration Program, when the Navy desired to establish a training program for calibration personnel, there were no comparable university or industry programs to use as guidelines. Commercial training was for the most part offered by various instrument and equipment manufacturers in the general areas of maintenance, repair, and calibration of their own particular equipment. While this type of factory training was, and is still, utilized widely by individual Navy activities, its limitations are obvious. One early approach taken by the Navy to fill the training gap was the use of on-the-job type training provided by experienced Navy personnel at higher echelon calibration laboratories. While this approach is still used occasionally, problems created by utilizing the valuable time of the technicians and engineers at the higher echelon laboratories, and also the fact that training is often provided on equipment of higher order than will normally be utilized by the trainee, precluded this type of training from satisfying overall program needs. It was in light of these shortcomings that the Navy initiated its own formal training program in electrical measurements during April 1959.

2. Early Development

The difficulties involved in convening the first formal training session were considerable. As a matter of fact, it was less than one week before the convening date of the course that the contract for the instructional staff was finally let. We were indeed fortunate that the instructors had been willing to devote a considerable amount of their own time to develop the course and training materials and, at the same time, familiarize themselves with the particular standards and associated equipment utilized by the Navy and also the Navy calibration procedures around which the laboratory experiments were developed. However, the lack of planning time did have some effects. For example, the laboratory experiments in capacitance and inductance measurements were scheduled for the first week of the course, whereas the pertinent lecture material was not provided until the third week. As another example, quite the opposite of the usual sequence, microwave measurements theory was presented before d-c and low frequency a-c measurements. Another problem which had to be overcome was one of morale. The Navy electronic technicians who attended the first course had received orders transferring them from present duty stations to Pomona, California for training, and then to a new ship where an electronics calibration facility was to be established. These men had no knowledge of the Navy Calibration Program, little awareness of the need for calibration, and

---

*Head, Field Operations Division, Metrology Department, Bureau of Naval Weapons, Pomona, Cal.

additionally, being higher-level petty officers with years of electronics repair experience, did not believe they required additional training. So, in addition to teaching theory and practice of electrical measurements, we had to do a selling job, that is, sell calibration to the trainees.

3. Objectives

The objectives of the Navy calibration training program may be stated as follows:

a. To develop and instill in Navy calibration personnel a progressive philosophy of the calibration problem.
b. To extend the capabilities of these personnel by increasing their knowledge of theory and techniques of measurements.
c. To instill in these personnel the respect for accuracy and precision in measurements.
d. To stimulate a continuing intellectual curiosity and a desire for scientific knowledge.

To accomplish these objectives, the lecture material is taught in terms of basic electrical measurement fundamentals and principles. The laboratory experiments familiarize the trainees with the standards and associated equipment encountered in the calibration laboratory, and also serve to introduce the calibration technicians to calibration data-taking, analysis, and reporting. Of special note, the Navy has always deliberately avoided teaching trainees to calibrate particular instruments, as this is, essentially, building in obsolescence.

4. Electrical Measurements Course Content

Calibration training, as initially established, was provided in two three-week sessions. One session covered d-c and low frequency a-c measurements in the areas of resistance, voltage, current, capacitance, and inductance; the other session covered coaxial and microwave measurements in the areas of attenuation, frequency, impedance, VSWR, and power. A large measurement spectrum utilizing electronic instruments, for example, VTVM's, oscilloscopes, signal sources, power supplies, and amplifiers, had been deliberately omitted as it was the general feeling at the time that most of the Navy technicians were reasonably familiar with this equipment. However, after the training program had been in operation for over a year, it was determined that this confidence was not completely warranted, especially with respect to calibration. Accordingly, the training program was expanded to eight weeks by adding a two-week electronics section to the three-week d-c and low frequency a-c course. The present electrical measurements course thus covers the entire frequency spectrum from d-c to microwave. As a retraining measure, special two-week courses in electronic instruments were provided those students who had taken only the original three or six weeks. A brief syllabus describing the general content of the electrical measurements course is appended to this paper.

5. Method

The teaching method established in early 1959 was for each of two instructors to have half of the group of trainees for the entire course. Part of each day would be spent in lecture and the other part in the laboratory, which consisted of one setup each of four or five experiments. Since it developed that certain instructors were more at home with the laboratory equipment than others, and also that one could cover certain lecture material more proficiently than another, the teaching method has evolved to its present system of having the entire class in lecture at the same time with the instructors shuttling in and out on schedule to cover specific lecture material. Duplicate setups of all laboratory experiments were established to handle the increased size of the laboratory class.

6. New Training Courses

The need for training of calibration technicians in measurement areas other than electrical led to the development of two additional courses which were presented for the first time this past spring. One of these is a four-week course in electro-mechanical measurements covering pressure, acceleration, force, torque, flow, mass, viscosity, and temperature measurements. The second is a three-week course in linear measurements, surface characteristics, and optical instruments. A somewhat less theoretical approach is taken in both of these courses than for electrical measurements. This is especially true for the linear and optics course in that the trainees are selected from laboratories thrice removed from NBS, at the level where the theory of the measurement by optical means, for example, is not as significant as the ability to make the measurement. In the electro-mechanical measurements course, although the basic physics of each measurement phenomenon discussed are adequately covered, the theory of the various electrical and electronic readout devices utilized is not of major concern, only their operation. A description of each of these courses is also appended.

7. Conclusion

One of the most salient features of the Navy's calibration training program is the emphasis placed on continual updating of the material covered and of the equipment utilized in the laboratory portion of the training. The continual review and assessment of new measurement and accuracy requirements in the field leads to revision of measurement systems and techniques at all echelons of calibration. This requires the Navy to readjust its calibration training to assure that the trainees are able to cope with new situations and are familiar with the latest measurement standards. A very recent example is in the area of frequency standards. Many of the Navy's calibration facilities are acquiring VLF systems in addition to the WWV system in order to establish a standard of frequency in such overseas areas as Japan, Guam, Scotland, and Spain, where WWV reception is poor or nonexistent. Comparison theory and techniques for the VLF system as well as WWV must be taught and demonstrated. Additionally, state-of-the-art breakthroughs, such as distribution of commercially manufactured coaxial thermal voltmeters to all Navy standards and calibration laboratories generates revisions to portions of the course material. The development of the linear measurements and optics training discussed above is a direct result of the requirements for calibration of optical instruments used in the navigation and firecontrol systems for the POLARIS fleet ballistic missile weapon system. It is believed that the Navy's ability to be flexible in its training of calibration personnel, to change existing curriculum, develop new· courses, and retrain, based upon actual fleet measurement and calibration requirements, aids in providing maximum calibration support to fleet test and weapon system equipment, which is, of course, our ultimate goal.

Electrical Measurements

This outline is intended to give a general picture of the course coverage. Exact content and order of presentation may be changed slightly as the course is developed. The course is divided into two sections, Section I on DC, Low Frequency AC, and Electronic Measurements, and Section II on Microwave and Coaxial Measurements. Section I will last five weeks, and Section II - three weeks.

Section I. D-C and Low Frequency A-C Measurements

1. Introduction. Measurement Philosophy.
    - brief history of measurement standards
    - present crisis in precision measurement
    - metrology - role and importance of precise measurement
    - unit systems - cgs, mks - dimensional analysis
    - experimental procedures defining units - international, absolute
    - standards - traceability
    - errors - significant figures - tolerances - interpolation
    - measurement concepts: direct, substitution, transfer
2. D-C Measurements
    - a. Theory
        - fundamental concepts
        - Ohm's Law - Watt's Equations
        - networks - Kirchoff's Laws
        - networks - Thevenin's Theorem
        - potentiometers
        - Wheatstone Bridge
        - the galvanometer
        - algebra through simultaneous equations
    - b. Instrumentation and Laboratory
        - resistance standards
        - standard cells
        - Wheatstone Bridge
        - Kelvin Bridge
        - Megohm Bridge
        - precision measurements of resistance in different ranges
        - precision measurements of potential
3. A-C Measurements
    - a. Theory
        - inductance
        - capacitance
        - need for the vector concept
        - a-c concepts - frequency, amplitude, effective (rms) values, phase, lead and lag, phasors
        - circuit concepts - reactance, impedance, time constant, RL circuits, RC circuits, RLC circuits, resonance, "Q" oscillators
        - a-c networks - Thevenin's Theorem, inductance bridge, capacitance bridge, "frequency bridge"
        - bridges - Hay, Maxwell, Owen, Schering
        - a-c meters
        - algebra and trigonometry through vectors, polar and j notation
    - b. Instrumentation and Laboratory
        - standard inductors
        - standard capacitors
        - inductance bridges - Owen (modified)
        - capacitance bridges - Schering
        - measurement of inductance and capacitance by bridges in different ranges
        - calibration of a-c and d-c meters
        - time interval measurement in microsecond range
4. Examination on D-C and Low Frequency A-C Measurements
    - theory
    - laboratory
5. Introduction to Electronic Instruments and Measurements
    - electronic instruments and their uses
6. Theory
    - a. Frequency Response Function and its Relationship to Electronic Systems
        - general definitions of frequency response
        - 3 db points
        - bandwidth
        - Q-bandwidth - 3 db point relationships
        - flat systems
    - b. Transient Response Functions
        - rise time
        - overshoot
        - ringing and smear
        - rise time - 3 db point relationships
    - c. Feedback Systems and Selective Systems
        - gain reduction
        - stability
        - LC systems
        - filter systems
    - d. Passive Elements
        - attenuators
        - loading
        - impedance matching
        - terminations
        - cables
    - e. Signal Sources
        - constant current
        - constant voltage
        - output and input impedance
        - maximum power
        - regulation
    - f. Audio Sources
        - Wein Bridge
        - beat frequency
        - calibration accuracy
    - g. RF Sources
        - XTAL types
        - LC types
        - stability
        - loading effects
        - calibration effects
    - h. Non-Sinusoidal Sources
        - pulse generators
        - square wave generators
        - VLF synthesizers
        - terminations
        - duty cycle
        - calibration requirements
    - i. Electronic Voltmeters
        - balanced amplifiers

        - stability
        - calibration factor
      - rectifier types
        - rectification factors
        - waveform error
        - turnover error
        - duty cycle error
        - source impedance errors
  - j. Wave Analyzers (selective voltmeter)
  - k. Oscilloscopes
    - time bases
    - video amplifier
    - dual beam transfer operation
  - ℓ. Frequency Counters
    - binary
    - heterodyne frequency meters
    - oscilloscope measurements
  - m. Time and Frequency Standards
- 7. Laboratory Practice
- 8. Examination on Electronic Instruments and Measurements
  - theory
  - laboratory

Section II. Microwave and Coaxial Measurements

- 1. Introduction
  - microwave concepts
  - typical microwave system
  - power, attenuation, VSWR
  - attainable accuracies
- 2. Transmission Lines
  - a. Theory
    - a-c source as a radiating system
    - two-wire transmission lines
    - infinite line - characteristic impedance
    - loads: matched impedance - short - arbitrary impedance
    - standing waves - VSWR, reflection coefficient, resonant sections
    - power in a transmission line
    - coax lines
    - waveguides
    - algebra through powers, roots, logarithms, decibels
    - Smith Chart
  - b. Instrumentation and Laboratory
    - slotted line; attenuator; sliding load; directional coupler
    - measurements in different frequency ranges of wavelength (frequency)
    - VSWR - reflection coefficient
    - attenuation
    - directional coupling
- 3. Attenuation
  - a. Theory
    - attenuation and impedance
    - insertion loss
    - mismatch loss
    - power ratios and decibels
    - more attention to algebra if necessary
  - b. Instrumentation and Laboratory
    - standard attenuators; directional couplers
    - square-law detectors; probes
    - measurements in different frequency ranges by insertion loss
      - method (low accuracy)
      - substitution method (high accuracy)
- 4. Frequency
  - a. Theory
    - microwave spectrum
    - oscillators
      - klystron
      - magnetron
    - harmonics-heterodyning
    - wavemeters - cavity resonance
    - counters - frequency dividers
    - more logarithms and ratios if necessary
  - b. Instrumentation and Laboratory
    - measurements in different frequency ranges
      - wavemeters
      - counter alone
      - counter with heterodyne oscillator (frequency converter)
      - counter with external oscillator and mixer
- 5. Power
  - a. Theory
    - power; frequency, amplitude
    - equivalent heating as a measure of power
    - calorimetry method
    - bolometers
      - thermistor
      - baretter
    - bolometer bridge
  - b. Instrumentation
    - thermistors and baretters; mounts
    - audio and rf substitution
    - measurements in different frequency ranges
      - power meter
      - substitution
- 6. Examination on Microwave and Coaxial Measurements
  - theory
  - laboratory

Electrical-Mechanical Measurements

1. Philosophy of Measurements
   - The significance of measurements
     - history of measurements
     - legal status of standards of measurement
     - systems of units
     - standards--accuracy and precision
2. Basic Physics of Standards and Accuracy of Measurement
   - Length, mass, time, temperature
   - Displacement and force
   - Time and frequency
   - Temperature
   - Electrical units
   - Pressure
   - Accuracy, precision, and error
   - Error classification and treatment
3. The Generalized Measurement System
   - What is a measurement (properties of materials)
   - Fundamental methods of measurement
   - The generalized system
   - Calibration
   - Sensitivity--period
   - Damping
   - Dynamics of a generalized system
4. Electrical Measurements in Mechanics
   - Introduction
   - Fundamental concepts
     - conductors
     - insulators
     - electric and magnetic fields
     - electric current
     - electric potential
     - resistance
     - Ohm's Law
     - power and energy
     - generation of emf
   - Circuits, d-c
     - series
     - parallel
     - voltage dividers
     - attenuators
     - bridges
       - balanced
       - unbalanced
     - potentiometers
   - Circuits, a-c
     - inductance
     - capacitance
     - impedance (a-c)
   - Electronics
     - amplifiers
     - VTVM
     - frequency measuring equipment
     - time measuring equipment (counters)
   - Transducers for basic mechanical measurements
     - strain gages
       - SR4
       - magnetic
       - differential transformer
   - Basic detector--transducer elements
     - preliminary comments
     - loading of the signal source
     - the secondary transducer
     - first stage devices
     - mechanical members as primary detectors
     - electrical transducer elements
     - advantages of electrical-system elements
     - variable resistance transducer elements
     - sliding contact devices
       - resistance potentiometers
         - potentiometer resolution
         - potentiometer linearity
         - potentiometer noise
           - loading noise
           - shorting noise
           - resolution noise
           - generated noise
           - high velocity noise
           - resistance strain gages
           - thermistors
           - thermocouples
         - variable inductance transducer elements
           - inductance
           - mutual inductance
           - reluctance
           - permeance
           - permeability
           - simple inductance transducers
           - two-coil mutual inductance arrangement
         - differential transformer
         - variable reluctance transducer
         - capacitive transducers
           - charging dielectric constant
           - charging area
           - charging distance
         - piezoelectric effect
         - photoelectric transducers
         - electronic transducer element
5. Intermediate Systems
   - Mechanical systems: inherent problems
   - Kinematic linearity
   - Mechanical amplification
   - Reflected functional amplification
   - Reflected inertial amplification
   - Amplification of backlash and elastic deformation
   - Tolerance problems
   - Temperature problems
   - Limiting temperature errors
   - Electrical methods
     - input circuitry
     - simple current-sensing circuit
     - ballast circuit
     - the voltage dividing potentiometer circuit
     - the voltage balancing potentiometer circuit
     - resistance bridges
       - Wheatstone d-c
       - Wheatstone a-c
     - impedance bridges
     - resonant circuits
     - amplification or gain
       - amplifiers
         - voltage
         - current
         - a-c
         - carrier
         - tuned
         - cathode follower
       - special circuits

      - RC-LC
      - differential and integration
      - filtering
      - coupling (impedance matching)

6. Terminating Devices and Methods
   - Introduction
   - Meter indicators
   - VTVM
   - Counters
   - Oscilloscopes
   - Oscillographs
   - Galvanometers
7. Measurement of Force and Torque
   - Mass weight - Force
   - Measuring methods
   - Mechanical weighing systems
      - equal arm balances
      - unequal arm balances
      - analytical balance
      - multiple lever systems
   - Elastic transducers
      - proving rings
      - load cells
   - Torque measurement
8. Measurement of Pressure
   - Introduction
   - Pressure measuring systems
   - Pressure measuring transducers
      - gravitational types
      - direct acting elastic types
      - indirect acting elastic types
   - Secondary transducers
   - Pressure cells
   - Measurement of high pressure
   - Measurement of low pressure
      - McLeod vacuum gages
      - Pirani gages
      - other gages
   - Dynamic characteristics of pressure measuring systems
9. Temperature Measurement
   - Introduction
   - Bi-materials
      - liquid in glass
         - calibration and stem correction
      - bi-metal elements
   - Pressure thermometers
      - liquid
      - gas
   - Thermo electric elements
   - Thermo resistive elements
      - resistance wire
      - thermistors
   - Measurement of resistance
      - Wheatstone bridge
      - Mueller bridge
   - Thermocouples
      - laws for thermocouples
      - thermocouple materials
   - Measurement of thermal emf
      - galvanometer
      - potentiometer
   - Pyrometers
      - total radiation
      - optical
   - Special problems
      - errors due to conduction and radiation
   - Calibration of temperature measuring devices
10. Measurement of Flow
    - Introduction
    - Primary or quantity methods
    - Rate devices
    - Velocity probes
    - Special methods
       - turbine
       - thermal
       - magnetic
       - sonic
       - mass
    - Flow characteristics
    - Obstruction meters
       - Venturi
       - flow nozzle
       - orifice
    - Variable area meter
    - Measurement of velocity
    - Pressure probes
       - total pressure
       - static pressure
       - direction sensing probes
       - special flow measuring methods and devices
    - Viscosity
    - Specific Gravity
11. Vibration and Acceleration Measurement
    - Introduction
    - Vibration meters and detectors
    - Accelerometers
    - Seismic instruments
       - vibration
       - acceleration
    - Calibration
       - static
       - dynamic
    - Natural frequency and damping
    - Response of seismic instruments to transients
    - Measure of velocity (angular and linear)
    - Exciter systems
       - mechanical
       - electrical
    - Vibration test methods

Linear Measurements and Optics Course

This outline is intended to give a general picture of the course coverage. Exact content and order of presentation may be changed slightly as the course is presented. The course will last for three weeks.

The basic objective of the course is to extend the capabilities of laboratory personnel by increasing their knowledge of theory and technique of mechanical measurements.

1. Introduction--Measurement Philosophy
    - brief history of measurement standards
    - present crisis in precision measurement
    - metrology--role and importance of precise measurement
    - unit systems--cgs, mks--dimensional analysis
    - experimental procedures defining units--international absolute
    - standards--traceability
    - errors--significant figures--tolerances--interpolation
    - measurement concepts: direct, substitution, transfer
2. Displacement measuring devices
    - low resolution devices
    - medium resolution devices
    - high resolution devices
    - super resolution devices
3. Optical methods
    - collimators
    - theodolites
    - transit square
    - jig transit
    - optical flats
    - polygons
    - lenses
    - prisms
    - beam splitters
    - interferometers
    - contour projectors
    - comparators
    - optical tooling
    - levels
4. Measurement of
    - length
    - planeness
    - parallelism
    - roundness
    - angles
    - screw threads
5. Surface finish
    - units of roughness
    - methods of analysis
6. Hardness testing

Session 6. Training of Measurement Personnel

Paper 6.7. Audiovisual Application to Equipment Calibration

J. C. Riley* and L. A. White*

This paper provides Audiovisual Training Aids for Calibration Laboratories.

1. Introduction

Why consider A-V technique for calibration training? Combining spoken material with pictures offers many benefits. Learning and supervision time are reduced, quality is higher and more uniform, and operators are useful at a higher level of skill.

We are going to see how to make an A-V program. There will be a discussion of the purpose of various programs and the methods of analyzing a procedure. We will look at techniques for easier and less time-consuming instruction writing and picture taking. Various ways will be shown to present the material for effective operator training.

2. Features of A-V Technique

There are many methods to train operators for calibration work. One of the best is the A-V technique. It can provide

- accurate effective training procedures
- standardized training at widely separated test facilities
- sharply reduced training time and cost
- increased performance for a given operator skill level
- availability of complete instructions for periodic training review
- less supervisor time required for training makes more of his time available for planning and deciding
- improved operator understanding and confidence

3. Making an A-V Program

3.1. The Purpose

The first thing to decide is the purpose of the training program. There are three areas to consider: teaching, practicing, and working.

Teaching requires an explanation of the procedures and their steps. The operator becomes familiar with your reasoning and gains the technical background to increase his ability.

Practicing requires test procedures to be done by the operator. Detailed explanations are minimized with the emphasis on developing speed and skill.

Working uses A-V technique for actual calibration tests being performed by trained personnel. Very little explanation is required. The emphasis is on optimum measurement speed.

These purposes can be modified by combining two of them into one program for a saving of time and effort. Tailoring a program has real benefits, but be careful not to try covering too many of the purposes with one program. If it is complete enough for teaching, it will probably be too slow for working.

3.2. The Methods of Analysis

Select a calibration procedure and study it. The design of the A-V program includes technical

---

*Electro Scientific Industries, Portland, Oreg.

analysis, methods analysis, and work simplification. Technical analysis looks at what the procedure does or will do and why. Methods analysis shows how to arrange the procedure. Work simplification may offer a better way to get the same results in less time.

These analysis methods tend to overlap and lose their identity. However, they must be considered either singly or together for a complete procedure study.

3.3. Building the Package

A lot of good material now needs to be put in order. The next step is to organize and present it for someone else to use.

Break down the procedure into a series of simple steps. Start by looking at the number of operations to be performed. Operations can be lead connections to equipment, pre-setting dials and knobs for a measurement, making a measurement, etc. Now see how many steps are needed for each operation. Use enough steps for a thorough explanation. Each step should contain a single instruction or idea.

Next, consider the pictures that will illustrate the operations. Do not use too many steps in a picture. Additional pictures with fewer steps will explain an operation with greater clarity.

A procedure breakdown can be done several ways. For example, a breakdown by operation can be followed with a picture breakdown. The steps in each operation are then fitted into the pictures.

The procedure now consists of detailed instructions and illustrations. Write or record how each instruction will be presented.

Figure 1 shows a good way to form clear, understandable instructions. This method of sentence construction will develop a style (see fig. 2) that helps the operator follow the instruction quickly and easily.

As the procedure develops it should have uniformity. It is poor practice to interrupt a style of presentation without good cause. Standardizing the instructions gives the operator confidence in the procedure and his ability to complete it. Full explanation of operational steps is given as they first appear in the procedure. If these steps or operations are repeated they can be shortened to speed up the procedure. However, they must still be adequately described. Most instructions should minimize explanations and come right to the point. Procedures for learning or for training will require greater explanation to be effective.

Illustrating a procedure cuts down the amount of spoken instruction needed. Many benefits of A-V come from using both the eyes and the ears to receive information. Therefore, pictures can and must add to a fuller understanding of the procedure.

Number the steps in a picture to save the operator hunting for the next one. This helps promote an even work pace. Relate the spoken steps to the illustrated steps with numbered arrows or markers. Color pictures will make the arrows or markers stand out for easy identification. An example is shown in figure 3.

A picture should show just the operation involved. Avoid distracting background that will cause the operators eyes to wander. For a detailed or intricate operation move the camera in. In figure 4 the picture becomes more important than any spoken instruction. It can describe at a glance an operation difficult to explain verbally. A picture should present the results of completing the steps presented with it. This shows the operator how the finished operation will look.

Pictures must be well taken to be effective. Avoid light reflection or heavy shadows in the work area. Proper focus is also very important. Setups with a depth of field will require special care to insure adequate focus.

3.4. Using the Package

The instructions are recorded and the pictures are made. How will we present this for the operators use? The most common form of A-V is the tape-slide package. The written material is recorded on tape. The pictures are 2x2 color 35mm slides. Commercial equipment is available to automatically play the tape using recorded signals to change the slides in sequence. Figure 5 shows an example of the equipment needed. This is the best way, but it is expensive. A separate tape machine and slide projector can also be used to give the procedure. This is less expensive but inconvenient to operate. Another way is to record the instructions on tape with the pictures in a ring binder or book. This is inexpensive and presents the procedure in a fairly useable form. The last way is to bind up the instructions and pictures into a book. Its disadvantage is that the operator must slow down to read the instructions. A procedure in complete book form or a procedure with tape instructions and picture book can be easily made into a full tape slide presentation at a later time.

4. Conclusion

Personnel training with A-V technique is worth considering for the calibration laboratory. Developing an effective A-V procedure requires a knowledge of the purpose, the analysis, making the procedure, and using it. The resultant procedure can be tailor-made to fit any circumstance and provide only those benefits desired.

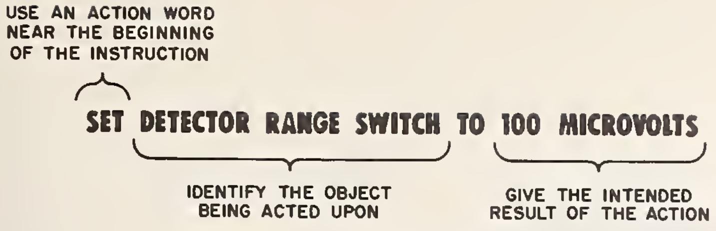
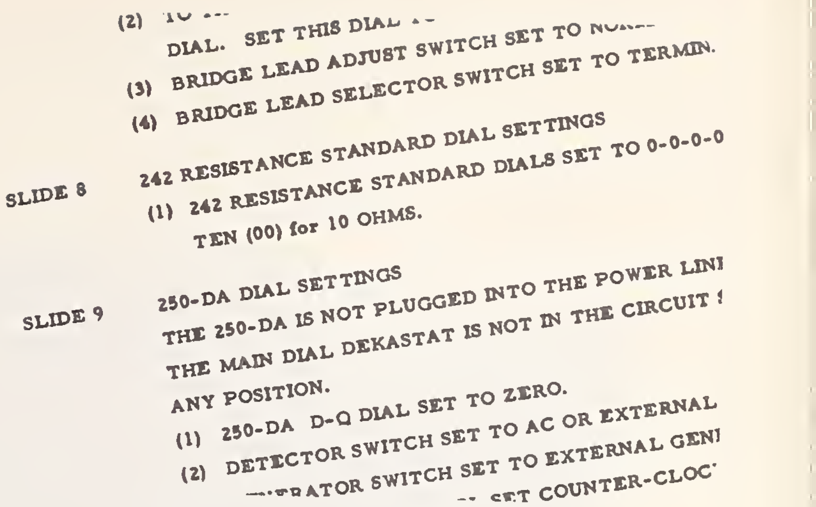


Figure 1. Instruction style.



(2) [illegible]

DIAL. SET THIS DIAL

(3) BRIDGE LEAD ADJUST SWITCH SET TO

(4) BRIDGE LEAD SELECTOR SWITCH SET TO TERMIN.

SLIDE 8 242 RESISTANCE STANDARD DIAL SETTINGS

(1) 242 RESISTANCE STANDARD DIALS SET TO 0-0-0-0

TEN (00) for 10 OHMS.

SLIDE 9 250-DA DIAL SETTINGS

THE 250-DA IS NOT PLUGGED INTO THE POWER LIN

THE MAIN DIAL DEKASTAT IS NOT IN THE CIRCUIT

ANY POSITION.

(1) 250-DA D-Q DIAL SET TO ZERO.

(2) DETECTOR SWITCH SET TO AC OR EXTERNAL

ATOR SWITCH SET TO EXTERNAL GEN

SET COUNTER-CLOC

Figure 2. Instruction sample.

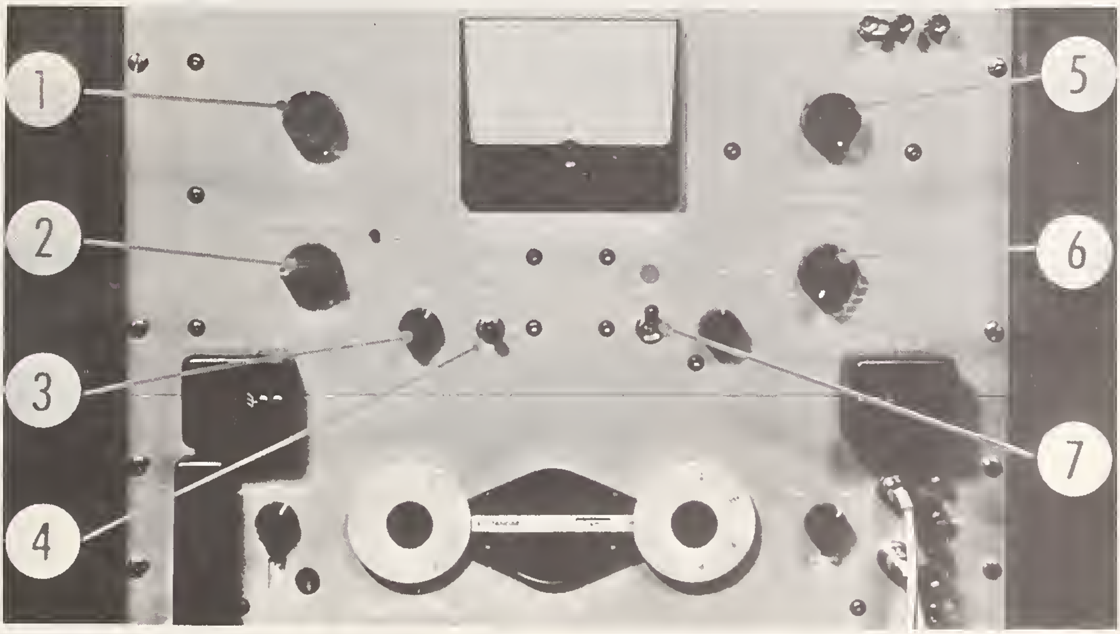


Figure 3. Illustration sample.

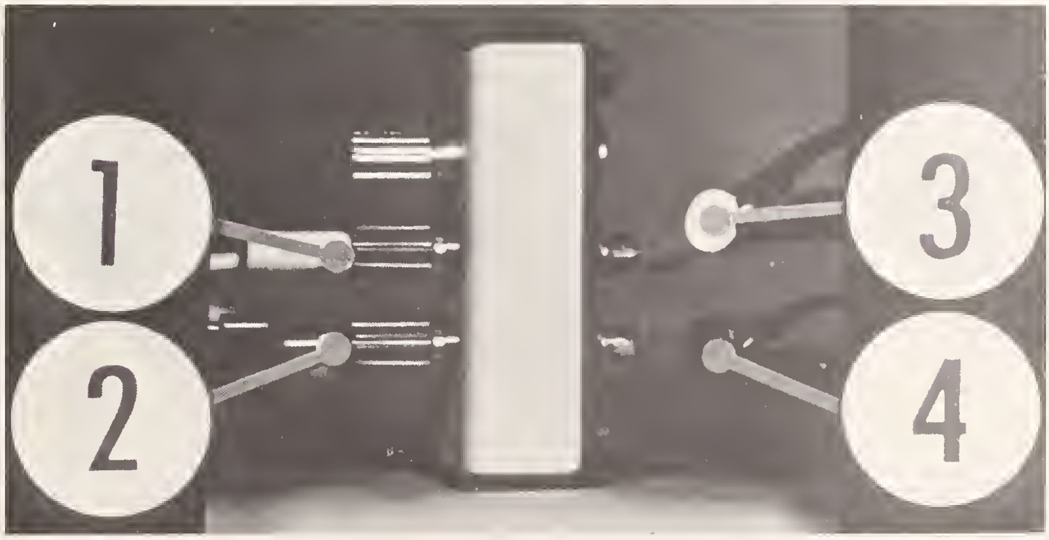


Figure 4. Illustrations for a small area.

Figure 5. Audio visual equipment.

Session 6. Training of Measurement Personnel

Paper 6.8. The Metrology Programs in the School of Engineering of The George Washington University

N. T. Grisamore*

This paper describes the Metrology courses at graduate and undergraduate levels which are taught at George Washington University, Washington, D.C.

1. Aims of the Educational Program

Prior to the establishment of the metrology program at the George Washington University, there was no full time degree-granting academic program in metrology. For this reason it was felt that students should be trained in general metrology without giving too much concern to specialization in any particular measurement area, e.g., electrical measurements, mechanical measurements, etc. Courses in specialized areas will be added as time and circumstances permit but it is felt that the greatest immediate gain will come about from the training of personnel in the area of general metrology. Graduates of this program will certainly be better equipped to go into metrology work than those from other technical programs. Most of what has been said above is meant to apply to the Technologist's and Bachelor's programs. The student will become more of a specialist as he proceeds through his graduate training but he must also retain an overall feeling for general metrology ideas.

As time goes on and a body of knowledge and experience is built up, both in the subject matter and in the educational programs of metrology, more specialization will, of necessity, occur. The present undergraduate curriculum does not allow for courses in medical, chemical, biological, nuclear, or atomic measurements and yet all of these are important topics in present day metrology. Numerous other areas of measurement can be thought of without too much effort, and the possibility of covering all of this material within a four or five year program becomes quite formidable. Eventually, it is hoped that academic programs will be devised to cover every aspect and facet of quantitative measurement including such topics as Econometrics and measurement aspects of Political Science and Law.

2. Implementation of the Programs

The course structure and content of the metrology programs are given under a subsequent heading. These courses were arrived at after discussions with a large group of people intimately connected with measurement problems and, of course, with educational programs. Fortunately, the university has received some industrial support in the form of equipment grants and money for this program. No new curricula, and very few of those which are long established, pay their own expenses through tuition charges: consequently the industrial support should be given considerable credit for the establishment and maintenance of the programs. At the same time, regardless of the excellence of any program, it is just a structure on paper unless students enroll in the curriculum, so some credit belongs to the students presently enrolled: thirteen undergraduate, six master's level, and two doctoral students.

---

*George Washington University, Washington, D.C.

3. Degree and Certificate Programs

The programs of study in metrology in the School of Engineering at the George Washington University range from Engineering Technologist to Doctor of Science. The first two years of study, consisting of 70 semester hours, are similar to those required of all engineering students, the only difference being in the substitution of three metrology courses, 12 semester hours, for other courses if the student desires the Engineering Technologist Certificate at the end of those two years. This degree corresponds to the usual two-year junior college degree. Most students plan to obtain this degree on the way to the Bachelor of Science degree. In fact, there are a few students in the regular engineering programs who plan to obtain this certificate by taking the above mentioned metrology courses in addition to their regular program.

The Bachelor of Science program consists of the two-year freshman and sophomore program mentioned above plus 70 semester hours in the last two years. There are 31 semester hours of metrology courses in the Bachelor of Science program. Courses in engineering, science, and mathematics, considered to be prerequisites to the metrology courses, constitute eighty-nine semester hours, and, in addition, twenty semester hours of work in the liberal arts are required. The metrology courses cover topics in statistics, design of equipment, instrumentation, standards, and precise measurements in the areas of heat, mechanics, and electricity.

The Master of Science program in metrology is the same as that in engineering. The student must, however, concentrate his studies and select his thesis topic in the area of metrology. At present there are obviously no students who could enter this program directly from an undergraduate program. Consequently, the students in this program take a portion of their work in the advanced undergraduate metrology courses. (Achievement over and above that of the undergraduate students is required of graduate students in these courses.) The nominal program of studies requires twenty-four semester hours of course work and a thesis. Students are not restricted to the twenty-four semester hours but may take as many hours as they desire within reasonable limits and provided that the additional courses have an intimate relation to the student's area of study and thesis topic. Course work required to make up undergraduate deficiencies in areas other than metrology cannot be applied to the program.

The Doctor of Science program is also the same as that in engineering. Programs of study are formulated by a faculty committee for each student based on the student's previous education and on a concentration of study in preparation of his thesis work. A reading knowledge of two foreign language is required. The specific languages are designated by the faculty committee and their selection is based on their importance to the student's area of study. There is no formal requirement regarding semester hours necessary for completion of the degree. The program is divided into two parts, a period of study leading to qualifying examinations and thesis work. It would be most unusual for a student to complete the program in less than two years after obtaining the Master's degree. A Master's degree in a scientific or technical field is required for admission to the program. In general, the program can be described as a period of tutorial study followed by work on a research problem.

In addition to these regular academic programs, a number of short intensified programs are offered in the summer months. These are usually one week in length and six to eight hours per day. A certificate is awarded at the end of these programs but they may not be applied for credit in any of the regular academic programs.

4. Metrology Courses

The courses in metrology may be separated into the following groups:

1. General Metrology

   Metrology 5-6; a two-semester, nine-hour course on the foundations of metrology covering standards, measurements, design of experiments, etc. This course is required in the Technologist's and Bachelor's programs.

   Metrology 201; a one-semester, three-hour graduate course similar to Metrology 5-6, primaily for the entering graduate student who has had no training in metrology.

2. Statistics

   Metrology 4; a one-semester, three-hour course in statistics as applied to metrology problems. Required in the Technologist's and Bachelor's programs.

   Metrology 203-4; a two-semester, six-hour graduate course in statistics and probability as applied to metrology problems. Metrology 4 or its equivalent is required as a prerequisite.

3. Instrumentation

   Metrology 113-4; a two-semester, five-hour course in instrumentation and transducers. Emphasis in these courses is on precision and accuracy of measurements. Required in the Bachelor's program.

4. Precise Physical Measurements

   Courses are given in a number of specific areas of physical measurements. The course descriptions are evident from the titles given below.

   Metrology 111-12; Precise Electrical Measurements; a two-semester, six-hour undergraduate course.

   Metrology 214; Microwave Measurements; a one-semester, three-hour graduate course.

Metrology 121; Precise Mechanical Measurements; a one-semester, four-hour undergraduate course.

Metrology 131; Precise Heat Measurements; a one-semester, four-hour undergraduate course.

Metrology 220; Precise Optical Measurements; a one-semester, three-hour graduate course.

A number of these courses are taught on the grounds of the National Bureau of Standards by personnel of NBS after regular working hours. The remainder are taught at the University also after regular working hours since most of the students are employed.

In addition to these courses two others closely allied to metrology are offered by the Department of Electrical Engineering. They are Electrical Engineering 111, Electrical Measurements, a one-semester, three-hour course required of all electrical engineering students and Electrical Engineering 225, Electronic Measurements in Psychometrics and Medicine, a one-semester, three-hour graduate course offered in conjunction with the university's School of Medicine.

5. Future Developments

Beginning in the near future the School of Engineering will be changed to a School of Engineering and Applied Science as a result of the increasing role of applied science in our present society. In addition to this change the method by which a student progresses through an undergraduate degree program will be considerably altered. The curricula will not be as rigidly prescribed as in the past. The most important effect of this freedom is that it will allow a student to progress as fast as he desires in a specific topic provided he can demonstrate that he has the prerequisite knowledge. The effect of this on the metrology programs will be to give students the opportunity to cover more of the specialized areas or to penetrate deeper into a single area. It is hoped that this procedure will allow students to make more efficient and valuable use of their time, not only in the metrology program, but also in all of the other programs. A second result of this change will be the requirement of courses in metrology in the curricula of all undergraduate students.

NBS Miscellaneous Publication 248
Proceedings of the 1962 Standards Laboratory Conference

Session 7. Calibration Recycle Analysis and Work Load Control

Paper 7.1. Instrument Recall Concepts and Policies

J. L. Hayes*

The importance of establishing sound concepts and policies to effect a high proficiency of laboratory operation and management is of major importance to standards laboratory organizations. Unless firm policies exist, the sizable investment in men and material in each standards laboratory will be wasted since the product the laboratory offers will be improperly utilized. This paper will discuss the different forms of instrument recall available to laboratories, handling and transportation policies, concepts and policies concerning equipment controls in total company operations, concepts concerning limits or restricted calibration of instruments, systems or "black-box" calibration policies, and the overall effect of these factors on both laboratory operations and customer or company operations. In each of the aforementioned areas of coverage, an analysis of statistics will be provided concerning both the present operations of over 100 companies as well as their planned practices in this area. These statistics will be based upon a special survey conducted by the NCSL Standards and Calibration Laboratory Work Load Control Committee.

1. Introduction

It is very well for most of us to proclaim the need for improved measurement standards, calibration facilities, and stronger management policies, but how can we be sure that these are actually producing results on the research bench, on the production line, at the test pad, and in the silo, or on the flight line? Only by an appropriate instrument control system for periodic recall can an activity be assured that the accuracy of all measuring instruments is maintained. It will be the purpose of this paper to instill in each of you a strong desire to establish effective recall concepts and policies which will enhance the quality and reliability of the end product. Standards and calibration laboratories are necessary and effective tools to the technical operations of the company and a vital economic investment that can achieve significant cost reductions and improved quality of the products manufactured.

Only through the establishment of sound concepts and policies will a high proficiency and effectiveness of laboratory operation and laboratory management occur. Unless these firm policies exist, the sizable investment in laboratory men and material will be wasted and the beneficial effects of calibration will be lost. Therefore, we have an ever-increasing responsibility to enlighten management concerning the effectiveness of calibration operations to achieve complete management support.

2. Desirable Instrument Recall Policies

The type of recall employed by a laboratory and the activities to which such recall applies is one of the keys to successful and effective laboratory operation. The types of recall available to laboratories are voluntary, advisory, and mandatory. While the latter sounds dictatorial in nature, it has been shown by the experience of all too many organizations that advisory and voluntary recall allows only part of the laboratory's job to be accomplished. When voluntary recall is used, the laboratory is at the mercy of every individual in the company, operates on a purely panic basis and is unable to plan for or substantiate input work loads. Further, it requires the ultimate in salesmanship on the part of the

---

*Metrology Technical Director, Bureau of Naval Weapons, Pomona, Calif.

laboratory manager to gain even a 20 percent degree of effectiveness. If everyone deeply believed in the value of calibration and were fully aware of its value, both voluntary and advisory recall would be fine. This utopian situation, however, is most unlikely to occur because of man's tendency to consider his work an island unto itself. Accordingly, his recognition of the need for measurement agreement between himself and thousands of others like him demands such professional and altruistic views as to make the actual application impractical.

Thus, we arrive at a realization that only through mandatory recall can measuring instruments receive required support. It then becomes important that we recognize the responsibility upon the laboratory manager that a mandatory recall policy imposes. Since mandatory recall is dictatorial, the laboratory manager must understand that he is cloaked with recall powers only so long as he exercises these in the framework of at least a benevolent dictatorship. As soon as he begins oppressing his customers with unilateral attitudes regarding mandatory recall, he has immediately sown the seeds which will destroy the essential policy. Therefore, mandatory recall demands a deep awareness of the needs of individual customers and the intelligent application of the technical problems involved. Thus, the recall of instruments which requires a research man to tear down a setup which it took him months to make and may take several weeks to get back on its feet is self-defeating. These situations must be dealt with individually and with a deep appreciation for individual requirements. We all know that we are not presently establishing recall intervals to such accuracies as to know that one day overdue on calibration on a three-month interval or one week overdue on a one year interval will really endanger the measurements that are made by the particular instrument. I suppose the best way to express the requirements of a mandatory recall system is that it be applied with both firmness and fairness to satisfy all of the factions that are affected by calibration--and that's just about everybody. Appendix A to this paper provides certain satistics pertaining to recall concepts and policies as derived from a recent survey conducted by the National Conference of Standards Laboratories Special Committee on Standards and Calibration Laboratory Work Load Control. Analysis of these statistics bears out the desirability of a mandatory recall policy insofar as laboratory management is concerned. Approximately 67 percent of the laboratories presently practice mandatory recall in some phase of their operation and 85 percent plan to practice it.

3. To Whom Should the Policies Apply?

Once a concept of mandatory recall has been accepted, the natural question arises as to whom the policy should apply. Very few of us would question the requirement that mandatory recall affects production and manufacturing operations. It is at this stage that final acceptance of the product is accomplished and where erroneous rejections and acceptances can take place unless test and measuring equipment is controlled through predetermined specification limits. We run into certain doubts, particularly on the part of those affected, when we enter either engineering or research and development programs, and, as anyone who has tried it knows, the R&D people are a far harder nut to crack than are those in engineering. What R&D and engineering people argue when confronted with the portent of mandatory recall of their instruments is that their measurements are relative (relative to what was done yesterday or last week, or something) and therefore do not require calibration support of an absolute nature. The argument against this stand is twofold. First, even if R&D and/or engineering measurements were purely relative, the normal accuracy decay of the majority of measuring instruments that might be used would necessitate recalibration to check the decay. If the data he is taking extends over any reasonable period of time, or if he wants to compare the results of what he did last year in an experiment with what he is doing this year, the improvements or changes he might have made to the experiment may well be lost through accuracy decay. Secondly, there is little, if any, R&D and engineering work done today which is an island unto itself and can be considered only relative. Such is particularly true in defense oriented activities where almost all effort is directed toward some application of the work at hand. Thus, if the results of R&D engineering are ever to become, either now or later, part of a usable product, the requirement for absolute measurement becomes mandatory. Otherwise, how could the original parameters and responses achieved in the research laboratory ever be made reality on the production line? This factor is all too often overlooked by R&D personnel and constitutes a significant problem to application engineers when they attempt to utilize the ideas and techniques generated in the R&D laboratory. One can't help but wonder whether the applications engineer really understands that many of his problems can be traced to the lack of absolute reference measurements, whether they be 0.1 percent or 10 percent in error, in the R&D laboratory.

Because of these reasons, it would seem inappropriate that any technical activity in an organization or company be exempt from a mandatory recall policy. Once again, it is the means of applying the policy and not the policy itself which determines a potentially deleterious effect upon R&D or engineering projects. Wisdom and sound judgment must continually be exercised by the laboratory manager in applying the mandatory policy to ensure that while he sets out to do good through calibration of the R&D/engineering instrument, he actually does harm to the rate of progress and the measurements being performed. For this reason, he must frequently establish a person-to-person relationship as well as a management system which allows reasonable degrees of freedom in applying the mandatory recall policy to R&D/engineering projects. Certainly, none of us wants to disturb delicate measurements in process nor tear down a process that has taken weeks to establish through the impersonal application of rules and regulations. We probably all have heard examples which embarrass the metrology profession such as microwave setups which were torn down to make a once a year VSWR check on a component while the engineer in charge of the project pulls out his hair over the thought that in one week he would have been through with that particular measurement. To compound the sin, he knows he will have to rebalance his whole system due to just the change in VSWR introduced due to the change in alignment by reintroducing the calibrated component back into his setup. These instances must be avoided by careful management controls, or such exceptional goofs will be broadcast by the injured party as the rule of the calibration group in the organization. The statistics given in appendix A bear out the desirability of mandatory recall for R&D and engineering groups. It can be seen that approximately 75 percent of the laboratories plan to incorporate mandatory recall for these two functions in their future practices.

4. Management Support to the Laboratory

There is a particularly interesting statistic which bears on the subject of adequate management support to the laboratory and the policies under which it must operate to be successful. It is noted that of those laboratories which indicated they planned to employ mandatory recall (85.4 percent), only 58 percent of the respondents indicated that a management policy was in effect which gave the laboratory authority for mandatory recall. One would assume that to have one of the elements, he would have to have the other. It would appear that many of the laboratories say they have mandatory recall but actually lack the power or the authority to enforce it properly. This leads one to the logical conclusion that a policy is only so good as the authority which supports it and that unless either the laboratory (or an agency within the company which the laboratory can technically direct) emforces the policy, it is little more than a piece of paper. There are many such policies existent in companies today which do not really state who is responsible for ensuring that the policy becomes effective and never becomes an active or vital part of operations. Many well written company documents are available which cover everything about the subject of calibration and the need and requirement for a mandatory recall system but they simply do not contain the necessary "punch line" which puts this system into effect. Accordingly, so long as company management pays only lip service to the concept and policy of mandatory recall and other laboratory support policies, they can be of little value.

5. Equipment Control and Handling Policies

The control of plant measuring equipment by proper means directly affects the ease and efficiency with which the laboratory operates. If the laboratory becomes aware of equipment in the company only by stumbling across it or seeing it when it is submitted for calibration, there is always that undercurrent of feeling that little more than 75 percent of equipment requiring calibration is actually being processed through the laboratory. We are all aware that plant equipment inventories are either inaccurate or have such confused administrative policies applied to them that many vital instruments never appear in an inventory. For this reason, many companies have found that the best means of knowing the location and status of measuring instruments is to serve as the company's central equipment control point in addition to the desirable role of being the central measurement intelligence center for the company. Because of the knowledge and awareness of laboratory personnel of the nature and application of measuring instruments, the laboratory becomes a logical choice for new equipment purchase review, recommendations for equivalent or more desirable instruments, recommending alternate instruments which have equivalent specifications but lower maintainability or out-of-service statistics, and recommendations and approval for more versatile instrumentation to do the same job at a lower price.

Another outgrowth of equipment control is the establishment of a loan pool of measuring instruments which serves a twofold purpose. First, the loan pool provides emergency support to new projects which could not meet progress schedules due to equipment procurement lead times. Second, the proper operation of a loan pool should significantly reduce the company's inventory of test equipment and concurrent capital outlay in measuring equipment through proper control of equipment when not in use. Many companies have experienced significant reductions in equipment inventories through proper utilization of loan pools. A by-product of the loan pool in terms of calibration is that with lower equipment inventories, lower costs of calibration result. Also, equipment in the loan pool which is not active does not require calibration at the schedule specified for in-use measuring instruments. This is accomplished by having a loan pool inventory for a certain type of oscilloscope of, say, ten instruments where, through experience, only four of these need to be kept in a fully calibrated condition for immediate issuance. The balance is kept on the shelf for unforeseen emergencies which allow sufficient lead time to pre-calibrate the instruments prior to use. Through this, the laboratory achieves higher efficiencies and lower operating expenses while maintaining the same quality and effectiveness in actual test and measuring operations throughout the company. In large companies, individual departments maintain loan pools; however, the advantages to both reduced equipment inventories and calibration hours are still achieved so long as consistent policies and the authority of the laboratory are maintained. A gratifying statistic in appendix A is that 80 percent of the laboratories presently possess either a laboratory or department loan pool and 83 percent plan one.

The policies established for the transportation and handling of equipment into and out of the laboratory have much to do with both laboratory efficiency and satisfaction of laboratory customers. Three methods of pickup and delivery are normally open to the laboratory; one by laboratory personnel, the other by user personnel, and the last by normal in-plant transportation personnel. Laboratories that have long logistic servicing chains sometimes utilize public carrier or user pickup and delivery. Where large laboratory operations are concerned, it has frequently been found most advantageous to use specifically assigned personnel for pickup and delivery of equipment requiring calibration. This is frequently accomplished by having personnel from the company equipment handling section assigned directly to the laboratory. The statistics of appendix A show a marked trend toward increased use of pickup and delivery by people assigned to the laboratory and significantly decreased use of normal in-plant transportation personnel.

6. Instrument and System Calibration Policies

Significant improvements in operating efficiency per instrument or calibration time per instrument can be achieved through the intelligent application of policies concerning limited calibration. As used herein, the word "limited" relates to applying calibration at special parameter points, frequency values, and accuracies vice calibration throughout complete manufacturer's specifications. Such a policy toward application of limited calibration entails the establishment and enforcement of a system to firmly determine requirements of equipment in using departments and to obtain necessary approval from cognizant inspection personnel. To achieve the necessary control to allay the fears of inspection personnel that limited calibrations are less desirable than full specification calibrations and to satisfy the needs of using personnel, the laboratory must serve as a central coordinating point. Many laboratories have found that their efficiency has increased significantly and the technical problems surrounding calibration have been reduced by the application of limited calibration. All too often a "blind" approach to calibration results in every instrument, no matter how weird its specifications, being subjected to complete calibration, or, at least complete calibration within the state-of-the-art or equipage limitations of the laboratory. The natural outgrowth of this results in requirements for voltage calibrations at 1000 Mc/s when actual requirements do not exceed 50 Mc/s or pressure calibrations to 100,000 psi when only 35,000 psi is the top pressure requirement in the company. While the desirability of the laboratory to continually improve capabilities is recognized, all too often priorities which are completely out of context with actual requirements place emphasis in the wrong areas and the laboratory finds itself with improved capabilities that are not required while less dramatic requirements are completely unsatisfied. For this reason, the laboratory has a deep and continuing responsibility to both apply limited calibration and to study company measurement requirements in consonance with a limited calibration policy and application system. It is recognized, however, that where no particular control is used or possible in the application of instruments, a limited calibration policy is difficult to establish and enforce. This is particularly true of those companies where only R&D activities are undertaken. Nonetheless, because of technical limitation of our calibration systems, it does become necessary that instruments be indicated as having received only limited calibration where state-of-the-art precludes any other determination. The statistics presented in appendix A, while not conclusive due to possible misinterpretation of the questions, indicate that either many laboratories are missing the boat in their responsibility to management to apply limited calibration or they have not firmed up opinions regarding just what should be done on this subject. At present, only 10 percent of the respondents indicate they have a systematic limited calibration policy in effect. It is the opinion of the writer that in future operations this figure should be much higher. Experience gained with numerous laboratories shows that less problems with inspection personnel and better service to customers at lower operating costs are the direct result of an intelligently applied limited calibration policy.

The subject of "system" calibration policies appears to be both one of the "oldest, yet the latest," things in the area of calibration. The decision to apply a policy whereby calibration of entire consoles is performed using simulated parameters versus the calibration of individual components comprising the console is not an easy one to make. The advantages of performing this kind of calibration are obvious so long as the calibration equipment which provides simulated parameters can be transported to the actual use area of the test console. Frequently, these consoles are so large that there is no other way to accomplish calibration as a system. Perhaps one of the impediments to date toward greater use of system calibration application is associated with the requirement for "in-place" calibration and the difficulty of configuring calibration consoles to accomplish this. The only alternative to many laboratories is to disassemble test consoles, take the components to the laboratory and perform calibration there. Unfortunately, when the console is reassembled, there is always that dark shadow of a doubt that the overall system comprised by the test console is actually operating to specifications derived from component calibrations. It is in this area that a combination of both limited and system calibration comes into strongest play. It is obvious that a great deal more study must be directed toward decision making processes in the application of system calibrations, and the intelligent and technically sound configuration of working level calibration consoles. The statistics of appendix A indicate that in this area, approximately 50 percent of the laboratories are performing some kind of system calibration but no one responded that the cost of this effort was too high. One cannot help but question whether the lack of answers given implies that little consideration has been given to this.

7. Why Bother to Calibrate ?

All of us engaged in this business of calibration know that it is an exciting and challenging profession and one to which we may become deeply devoted. It allows us readily to relate to and follow God's pattern of creating order out of chaos--precision out of imprecision. Perhaps that is why many of us are attracted to it. However, management may well not view the profession or its need in such dramatic terms nor see it as being really necessary or vital to the company's operations. No doubt each of us has been called upon from time to time to present dramatic proof to management that calibration will cure all technical ills and save all sorts of dollars. What is most frequently asked is to indicate what program fell short, did not meet production schedules, or what missiles failed because calibration was inadequate. Most of us are without such examples or, if we have examples, they are couched in such terms as to leave management completely unimpressed.

Unless management is shown the value of calibration, in both financial and improved reliability terms, a deaf ear is a natural result and the establishment and support of critical laboratory operating policies cannot become reality. This must become one of the early objectives of all of us and one to which our energies within NCSL should be devoted. I would propose that a collection be made of examples or treatises on the subject of the value of calibration and distributed to NCSL membership for emulation of calibration's value to management. Such "witnessing" could spell the difference between management's support for successful operations and a laboratory handicapped in accomplishing its mission. The delegation of management policy and management authority to calibration operations within a company are not something that will be handed to us on a platter; like respect, they must be earned and they must be recognized as being essential to the well-being and effectiveness of any standards or calibration laboratory.

Appendix A. Recall Concepts and Policies
Survey Summary

Surveyed subject	Laboratory responses	
	Present practice	Planned practice
A. Recall policies		
1. Voluntary	%	%
a. R&D	28.0	8.9
b. Eng	16.2	5.9
c. Prod/Mfg	10.3	2.9
2. Advisory		
a. R&D	28.0	17.7
b. Eng	35.2	17.7
c. Prod/Mfg	22.1	11.7
3. Mandatory		
a. R&D	44.0	73.4
b. Eng	48.6	76.4
c. Prod/Mfg	67.6	85.4
4. Authority in being for laboratory to recall mandatorily	55.0	58.0
B. Handling/transport policy (pickup and delivery of equipment):		
1. By personnel assigned laboratory	51.4	57.0
2. By normal plant personnel	40.0	29.9
3. By equipment user	35.6	31.4
4. By public carrier (external service only)	11.4	11.4
C. Plant instrument control policies		
1. Laboratory controls and loans all equipment to users	15.7	27.2
2. Laboratory maintains company loan pool	40.0	55.6
3. Individual depts maintain loan pools	38.2	32.8
4. Loan pools used either in laboratory or individual dept	80.0	83.0
D. Limited calibration concepts in effect		
1. Only when user requests	70.0	Inconclusive responses.
2. Used wherever possible through systematic review	10.0	Do.
3. Not at all	30.0	Do.
4. Equipment receiving limited calibration: temperature/pressure/force transducers, VTVM's, RF signal generators decade resistance boxes, recording galvonometers, digital voltmeters, indicating meters.		
E. System Calibration Policies/Concepts		
1. Applied to product/test/inspection consoles	51.9	53.4
2. Applied to working level calibration consoles	23.6	33.0
3. Cost too high to institute	0.0	0.0
F. Size of laboratory operations of respondees		
1. 10 or less people	30.0	
2. 11 to 25	21.4	
3. 26 to 50	11.4	
4. 51 to 100	12.9	
5. 101 to 300	1.4	
6. 301 and up	1.4	
G. Type of Laboratory Responding		
1. Facility "primary" laboratory only	15.5	
2. Working level calib. laboratory only	31.1	
3. Operations at all calibration levels of concepts (primary and working)	53.4	

Session 7. Calibration Recycle Analysis and Work Load Control

Paper 7.2. Work Flowing Planning

J. M. Aldrich*

A significant aspect in effective management of a standards or calibration facility is the proper utilization of manpower in terms of input work load and the knowledge of just what that work load will be. A major impediment to effective laboratory operation is that the laboratory supervisor is frequently unaware of new work load requirements being generated by other company operations such as R&D, engineering, manufacturing, and final inspection. A deeper appreciation by management of the need for work load forecasting and concomitant laboratory manpower requirements is essential if the standards laboratory is to operate in an efficient manner. This paper will concern itself with methods of ascertaining advanced or peak work load requirements from which appropriate work loads can be derived, the estimation and application of manpower requirements on the basis of "standard" hours per instrument and instrument category, and the appropriate engineering and clerical support required in terms of instrument volume processed. Particular emphasis will be placed upon the necessity for data feedback in terms of new product requirements as well as laboratory calibration operations in order that both advanced planning and "standard" hours can be ascertained and implemented. In each of the aforementioned areas of coverage, an analysis of statistics will be provided concerning both the present operations of over 100 companies and their planned practices in this area. These statistics will be based upon a special survey conducted by the NCSL Standards and Calibration Laboratory Work Load Control Committee. The possibility of NCSL making available a document containing the consensus of opinion on standard hours per instrument in terms of volume of instruments processed will be a significant summary of these statistics.

The subject of this session today, particularly of my part in this program, is the discussion of laboratory work flow planning. A significant aspect in any effective management of a standards or calibration facility is the proper utilization of manpower in terms of "input" work load, the knowledge of just what that work load will be. A major impediment to effective laboratory operation is that the laboratory supervisor is frequently unaware of new work load requirements being generated by other operations, such as research and development, engineering, manufacturing, and functional test inspection. Deeper appreciation by management of this need for work load forecasting and concomitant laboratory manpower utilization is the essential ingredient for the proper scheduling of any successful operation of a standards laboratory function.

This session will concern itself with suggested methods of ascertaining advance or "peak load" work requirements, from which appropriate work loads can be derived. This will be shown by the detailed description of the estimating phase and by the application of manpower requirements on the basis of standard hours per instrument and by instrument category. In addition, we will show appropriate engineering and clerical support required in terms of the instrument volume processed.

Particularly emphasized is the necessity for data feed-back terms of new product requirements, as well as laboratory calibration operations in order that both advanced planning and standard hours can be ascertained and implemented.

In addition to each of the mentioned areas of coverage, we will discuss in detail an analysis of the statistics returned by many of the participants at this session. These statistics comprise a summary of more than one hundred companies answering the questionnaires that were sent out previously.

The objective of this session is to analyze present practices and concepts in the advanced

---

*Formerly Executive Assistant to Director of Quality Control, Ryan Aeronautical Company, San Diego, Calif.

planning of calibration and repair of measuring and test equipment. We will not only study the answers that were received on the questionnaires as to what the current trends in present practices are, but we will also attempt to analyze what these companies are going to do in the form of advanced planning with regard to calibration.

Basically, this session will report on a study and the questionnaires submitted to this committee. These will be presented in three parts, which are the present practices, the planned practices, and a recommended practice. A study of the statistics revealed by the questionnaire in the field of advanced peak load planning showed that company laboratories are notified of new requirements prior to their ordering on an order of only 12.8 percent with a potential increase of only approximately 2 percent, or a total of 14.8 percent. In general, standards laboratories do not know in advance of new equipment requirements in 87.2 percent of those answering the questionnaire.

Under the question of laboratories being notified when new equipment is ordered, the following statistics were apparent from the questionnaire:

a. 38.6 percent of the companies answered that at this point they were notified.

b. 44.3 percent indicated they planned to initiate practices to be notified at this point.

At the present time, over 61.4 percent answering the questionnaire indicate no communication or knowledge of new equipment ordering.

The majority of the companies answering the questionnaire indicated that the first time they became aware of new equipment in the house was when it appeared for calibration. It seems logical, therefore, to deduce that many pieces of equipment are being used by companies without calibration and control, since over 40 percent have indicated that no information was received to control this category.

Statistics further reveal that in operations planning, those employing work load forecasting amounted to only 41.3 percent. Under the proposed plan, an increase in this category of work load forecasting will be 62.3 percent. Approximately 37.7 percent will have no plans in the future for work load forecasting. In manpower forecasting, there are currently 21.6 percent actively engaged in this type of forecast. Future planning indicates an increase to 35.1 percent. Actually, 64.9 percent had no plans for manpower forecasting. 22.9 percent of the total returns made comments on many different categories. They have been summarized into seven major headings. They deal with the fact that most people actually use empirical values or actual times. There were a number of companies which were too small to use sophisticated manpower or planning-forecasting methods. In some cases, there has been an expression of no management interest. A very small minority has indicated that standard hours have been developed and are now in use. Certain companies are using government standards, but others object to the use of these standards, feeling they are not realistic. It was interesting that there were many different viewpoints. For example, an exception was taken to the number of operations per instrument, rather than by instrument, to establish the standard hour value of calibrating an instrument.

Certain questions were misunderstood, as indicated by the replies. Perhaps the questionnaire was inadequate, but there were problems of professionalism versus good business practices. Apparently, some members indicated that they are professionals rather than everyday production personnel. These statistics, which were developed for management and for the managers of standards laboratories, are required for the plan of future industrial needs. These needs demand that we schedule and plan for effective practice of calibrating and repairing standard laboratory equipment. It would be a disservice to this session to digress at this point whether or not standards laboratory technicians have professional standing, or whether they are to be considered as part of the regular production task. This question was not in mind when their committee designed the questionnaire. It was, however, interesting that a great many people answering the questionnaire mentioned professionalism.

As indicated by the statistics, 40 percent of those answering the questionnaire indicated their willingness to supply NCSL Committee with standard hours developed by their organizations. Again, 40 percent indicated willingness to furnish volume records for use by the NCSL Committee on work load control. The NCSL could provide a service to all its members on manpower and planning forecasting. This service would be in the form of suggested methods utilized by other members. In order for NCSL to be of service to its members in the field of planning forecasting and manpower forecasting, cooperation of each of its members is required for any effective data analysis and distribution.

In the field of statistical information retrieval (a very complicated technique at the least), we need everybody's input, whenever questionnaires are sent to each of you. For instance, to give you an idea of the performance of the member groups, we find that only approximately 70 percent of those contacted have actually answered the questionnaire. Of this 70 percent, only 50 percent answered all the questions, 22 percent answered the comments, and only 25 percent answered the questions on proposed planning. With these differences in percentages, we can apply some reliability percentage theories. We find that only 15 percent of the questionnaires to NCSL members are valid for evaluation and upon which we can base a logical and reasonable analysis of what industry is doing in the field of calibration and repair and what they plan to do in the future. We need better member cooperation in the future to develop meaningful statistical information of value to each member.

Let us now consider that part of our subject in the field of recommended practice to help you develop your own work load plan. This approach was developed last January 25 by your committee at Boulder, Colorado. Basically, we are talking about a three-step approach: advanced planning, operations planning, and peak load planning. The first step, advanced planning, orients one to the new project requirements. Are there new measurement areas or techniques required? Increased tolerance requirements? What effect will this new project have on the work load? What are the critical timetables as far as scheduling is concerned? After evaluating the new project requirements in the field of measurement areas, we then evaluate needed facility requirements. What new types of equipment will be required? What new types of calibration standards and repair techniques will be involved? And, upon evaluation of this, exactly what level of project personnel will be required to operate the equipment requirements and the standards calibration requirement? The next step in line, obviously is, "Who's going to pay for it and how?" budget questions which are of prime consideration in any project. Often project monies available will dictate a further review of requirements and needs to establish absolutes. In operating any business for the benefit of company stockholders, evaluation must be on a "need" basis as to whether or not these are finite techniques of requirements, or whether they are just luxuries.

In the second phase, operations planning, once the new equipment has been evaluated it must be melded into current loads by establishing categories requiring calibrations, whether in the micro wave area, audio instrumentation, or other areas in which this new equipment falls. Once this is established, we can work on the calibration requirements of each of these new pieces of equipment. What is their level of service? The frequency and extent of servicing? Establishing the calibration requirements will then give us the quantity by category, by instrument, by operation, and according to a schedule that the new project will require.

The establishment of standard hours per category per instrument per operation will then give us the total hours requirements as set up by the schedule. We now have our total work load determination by hours and proceed to establish manpower categories--what is required by technician, by type, and by skill. Adjusting these manpower establishments and determinations according to the operating schedules provides the necessary information to do an adequate job in level loading. This means elimination of peaks and valleys inherent in any standards laboratory operation.

Once the schedule of manpower and equipment has been set up, we reach our most important phase of operations planning--performance monitoring. This is the reporting system, the corrective action, to level load where obvious inconsistencies have arisen. In this manner we get a second crack at what we had in the first place, where our information was limited to empirical theory in advance of the actual facts. We can now make adjustments by reporting, either through a mechanical means or a manual means, exactly how our manpower is working with regard to the equipment being calibrated. Once this performance monitoring is in effect, it provides an easy means for the third step in our recommended practices--peak load planning.

In the peak load operation planning we have methods of anticipation to consider. Obviously, the most sophisticated method of anticipation is pre-programming our information into a computing device and thereby arriving at what our peak loads will be; or by manually computing the daily requirements by scheduled operation of the calibration requirements of instruments. In addition, a factor for the repair and maintenance of this equipment must be added, as derived from past experience. This evaluation of work loads will then provide the necessary impact to supply us with information for remedial action. This remedial action will be a method by which we correct our original error as developed through the actual operation of the equipment. We can do this by either rescheduling to provide overtime for peak load, by subcontracting to calibration laboratories, hiring part time hires such as "moonlighters," or, finally, if the load justifies it, we can hire additional manpower to take care of our increased work load. Conversely, we should examine the aforementioned means to plan and execute "valley" conditions by reducing personnel to a level load condition and utilizing to the utmost subcontracting, occasional labor, etc., to take care of incidental increases in calibration loads.

In summary, we can streamline and customize our operation to most effectively utilize our manpower and do the job that is required in providing our industry with optimism and reliable service in the field of calibration and repair.

Each company has its own requirements by virtue of such factors as size and product. Good business sense places on each of us the responsibility to plan and control our standards laboratory operation in the most effective and economical operation possible.

At Ryan, we put into practice the IQ approach to implement those sound practices.

At Ryan, IQ means Insured Quality--the basic atmosphere necessary for the effective control of measuring and testing instrumentation.

Session 7. Calibration Recycle Analysis and Work Load Control

Paper 7.3. Operating A Calibration Laboratory

M. Rothbart*

The differing details of laboratory operations have a direct influence on the efficiency and reliability of the laboratory, and all too often these details are overlooked. The difference between a successful operation and a shaky operation depends on how instruments are processed once they arrive at the laboratory for servicing. These processing techniques bear directly on a reduction of standard hours per instrument and thereby influence every other aspect of laboratory operation and management. This paper will discuss the means of detecting incoming instrument servicing needs, the application of limited or restrictive calibration, the application of instrument specification derating, methods of delivery, shipping and receiving in plant and outside plant operations, factors influencing "in-place" versus "in-lab" calibrations, and methods of identifying and tagging instruments during the calibration operation. In each of the aforementioned areas of coverage, an analysis of statistics will be provided concerning the present operation of over 100 companies as well as their planned practices in this area. These statistics will be based upon a special survey conducted by the NCSL Standards and Calibration Laboratory Work Load Control Committee.

1. Introduction

The three major ingredients of a calibration laboratory are people, equipment, and facilities. Calibration laboratories are set up to provide service for a product-centered operation or support for military needs, or to serve industrial users.

Effective and economical service requires consideration of three basic steps, taken in sequence: the routing of work into the lab, through the lab, and back to the user. These functions are equally important and must be planned according to specific need.

The NCSL Workload Committee is concerned with these functions and has tried, through the Workload Committee special survey, to find out what some companies' problems are and in what direction people are going.

I will talk a little about the survey results and some of the things we feel are prime factors to consider in planning workload handling. The survey results indicate that there are no hard and fast rules for operating a calibration laboratory.

Processing instruments requires many decisions. What are the factors to weigh in making a decision about handling each instrument:

1. Economic--Is this instrument worth servicing? What is its history--reliability?
2. Schedules--How much time can be allotted per unit--at what point in time can we say we have a reliable instrument?
3. Equipment utilization--Line planning requires that laboratory instrument workloads be spread for maximum equipment utilization.
4. Predictability--History and projected usage is important.
5. Computer time--Are the records kept by a system used for other things with allocated time for calibration laboratory requirements?
6. Priorities--Self-explanatory--who gets this and how are they policed?
7. Personnel--Distribution and usage.
8. Planning function--Scheduling workload control.

---

*National Astro Laboratories, Inc., Pasadena, California.

These eight decisions have three common goals; traceability, accuracy, and reliability. Weighing all these factors in terms of each laboratory need brings us to the elements of equipment flow.

Incoming instrument screening needs, if our survey is correct, are generally detected as a part of routine calibration. This is fairly standard and works best where service and calibration are performed in the same facility. In facilities where the calibration laboratory is separated physically, a certain amount of time is lost when a high percentage of service work is expected. Eighty percent of the people surveyed indicated that screening was done as part of the calibration.

Servicing requirements can very often be detected by checking history records. We find in our laboratory that not only do the instruments have chronic peculiarities but so do the users. This history is helpful in maintaining parts inventory and knowing where to route for beginning processing. Here we might add that if standard hours are used, deviations due to pecularities should be monitored and noted on the record card. Successfully used, this technique will feed valuable information back to the user. This survey revealed that very few people are interested in changing the concept of detecting needs at the point of calibration.

The subject of limited or conditional calibration is a really grey one. Our survey indicates that over half the laboratories do limited calibration when specifically requested to. The latest official word on this subject is MIL C 45662A, which states that accuracies be compatible with the intended use. This is a calibration system requirement specification although it does apply to individual instruments. Twenty percent of the replies indicated that this was done in special bays and setups and half of those answering plan to continue this practice. Ten percent said that this method is used in specific cases, after checking with the user. About 50 percent of these people do not plan to continue this practice. I can sum up by saying that limited or conditional calibration is a subject to be better defined by NCSL as a help to all of us.

Instrument derating specifications, while not normally a part of a calibration procedure, must be considered for a number of reasons. Certain instruments are sold not meeting all published specifications. Some instruments, while they are marginal and meet specifications, will not maintain accuracy over a reasonable period. There are instruments which meet published specifications, but, in a particular environment or under certain usage conditions, will not consistently meet specifications. For the reasons listed, derating is necessary and should be a paperwork process preceding actual calibration. Again, back to the survey. 54 percent feel that an instrument should be derated when it does not meet advertised specifications; 27 percent additionally planned to incorporate this philosophy. 48 percent would derate if the calibration system was "state-of-the-art" with a limited confidence ratio. 25 percent would downgrade with consent of user and inspection; another 17 percent planned to establish this practice. There were only three people who would derate without concurrence of user and inspection.

There is a tendency on the part of some electronic instrument manufacturers to write specifications before designing instruments, and then to marginally meet specifications with a few handmade models, without considering that production tolerances degrade accuracy in production runs. Very often, derating on the part of the user will force the manufacturer to get on the ball and clean up his gear.

Shipping and receiving of scientific instruments comes in for some interesting comments. Apparently all of us are aware of the problem of handling instruments. Over half of the people commenting on this point in the survey used some special method of shipment. Roughly one-third used padded or shock-mounted trucks, carts, etc. One-third used special containers and one-third used special personnel. All indicated that they planned to improve in this area. It is a problem and it is recognized. We have used all three methods and each has been successful. It seems as if everyone in the calibration industry is aware of the problem. Much work needs to be done on an approved-type container. I think that this might be an excellent project for some NCSL committee. We have designed two models, different sizes, that seem to work out very well.

The subject of instrument down time is always an interesting one. From our standpoint we must assume that whatever the case may be, the user needs all of his instruments right now. Knowing this, we don't have to worry about getting them back on time, it's just a question of how long too long is. I feel that it is a user problem to plan system down time, including individual instruments, to determine maximum utilization requirements. I don't see how a calibration laboratory can concern itself with planning user's system down time.

Our survey indicates that 42 percent of the companies questioned manage a 3-day down time, and that 34 percent were working toward this goal. 30 percent turned around in two weeks, 24 percent one week and 14 percent managed ten days. It is significant here that 34 percent were working toward three days while the balance, on a longer down time, did not plan to shorten up. Here I feel that the size of the company is a significant factor and we did not try to evaluate turnaround based on relative sizes of operations.

The "In-Place" versus "In-Laboratory" controversy still goes on and here 75 percent of the people do in-place calibrations only on instruments that cannot be brought to the laboratory. Very few people feel that enough concentration of instruments exists in the field to justify taking the laboratory to the instruments.

Where in-place calibration is performed, normal laboratory equipment is transported on carts. This applies in 40 percent of the cases where remote calibration is performed.

A-c and d-c power supply calibration accounted for about 50 percent of the transportable calibration consoles, 20 percent report doing VTVM's this way, 35 percent did panel meters remotely, and 12 percent did scopes and signal generators remotely. 35 percent do recorders in place. The planned practice in all cases is away from on-site calibration.

The criteria for deciding when to engineer and build, versus commercial purchase, requires so many considerations that no attempt is made here to present significant data. Value analysis is an individual study. Critical need is the reason given for building in-place calibration gear. Engineering intuition seems to be the basis for deciding what to build in the way of special calibration gear.

Instrument calibration tagging seems to be an area where 97 percent of the people agree. Tagging is a must; what to put on the tag is another story. The confidence level of a tag is generally acceptable. Calibration tags are used by 97 percent of the people surveyed. 56 percent used reject tags where appropriate. 52 percent used the limited-use tag. Special data on limited use is included by $1\frac{1}{2}$ percent. Date calibrated and date due are used by 60 percent. MIL C-45662A clarifies the tagging situation by stating that calibration date be included, NR 520 says only due date is necessary. NASA 200-2 requires that calibrated instruments be maintained on variable maintained data. It seems to me that one area of activity for an NCSL committee would be to work on standardizing these tags--size, color, and date.

In presenting this information, we realize that there is much to do. While the data is incomplete, it is significant and certainly gives us guidelines to work from.

Session 7. Calibration Recycle Analysis and Work Load Control

Paper 7.4. Instrument Recall and Recycling Analysis Through Automatic Data Processing Methods

J. R. Van de Houten*

Once a company's standards and calibration organization begins to implement rules for the calibration of equipment and opens its doors for operations, it finds itself investigating, designing, and establishing a series of administrative controls and techniques to monitor, simplify and enforce these rules, and assure proper laboratory operation. One of the most important aspects concerning, and sometimes plaguing, the standards laboratory is just how to get test equipment into the laboratory from using activities and how often it should be calibrated to maintain reasonable reliability levels. While this paper does not endeavor to resolve the problem of how often to bring the instrument to the laboratory, it will attempt to move one step closer to obtaining valid information from which proper technical decisions can be made. Virtually the first item that a calibration activity must establish is a recall system to obtain the instruments and a data collection and analysis system to ascertain when instruments are due for recalibration. These systems range from simple manual methods, often with locally designed forms and reproduced by ditto, multilith, or other such processes, to relatively refined automatic systems using magnetic tape, computer techniques, and general ADP (automatic data processing) methods. Although there are many manual systems now in use, there is a strong trend toward ADP to improve laboratory operations. From such ADP techniques, considerable products and by-products emerge which do not only bear fruit for administrative decisions such as standard hours, repair time, etc., but provide the basis for technical decisions concerning recalibration cycles and general reliability levels of test equipment owned by the company.

Introduction

As we are all deeply involved in the business of making measurements, most of us are probably familiar with, and possibly even have posted on our office walls, Lord Kelvin's remark:

> "I often say that when you can measure what you are speaking about and express it in numbers you know something about it; but when you cannot measure it, when you cannot express it in numbers, your knowledge is of a meagre and unsatisfactory kind."

Actually, this statement is valid not only with respect to the science of metrology but also in the management and control of our standards and calibration operations--the area in which the Workload Control Committee's efforts are directed. It is necessary to measure, by some method, output, efficiency, effectiveness, direct and indirect costs, and many other parallel or subsidiary characteristics of our organizations.

It is also desirable to have some measure of what similar organizations are doing in these areas--no matter how subjective these measures may be.

We are all employed in an area that has received a tremendous impetus throughout the country in the last 5 to 10 years. As a result, I'm sure most of us have been called "empire builders" at various times, and, in a sense, we probably are. Yet, the success of our organizations and our future as individuals depends far less on our ability to build empires than on our ability to operate an effective and efficient operation--and to provide our management with: first, sufficient information for them to determine the efficaciousness of our organizations; second, sufficient justification for an affirmative decision in those areas where we must be given increased funds, authority, or backing; and, third, information through which they can establish overall objectives for our organizations. Besides these

---

*Aerojet General Corporation, Sacramento, Calif.

reasons, I often get the impression from the people we service that they are more concerned with how well we handle our work load than on the technical adequacy of the service.

This paper is one of four presentations based on the data returned on the survey form distributed this spring by the NCSL Workload Control Committee. This report covers the data in Section II, Recall System and Methods, and Section V, Data Collection and Utilization Methods. The data received does not, of course, have the parts-per-million accuracies many of us deal with in the laboratory. However, in spite of some limitations, this survey has provided the best measure available today of the techniques used by various companies for controlling their operations. Since I obtained this data before the rest of you and had to study it to prepare this presentation, I have already put some of it to work in my department. I have been considering its use in a number of other areas, and I believe it will be of considerable benefit to anyone who rationally applies these measurements for the purposes I've already mentioned.

Later I have indicated the actual data returned as factually as possible, although, in a few cases, some interpretation was necessary. I have tried to show correlations between the answers given in different sections of the questionnaire and to break down in various ways the data returned where it appeared significant. In several cases I have provided additional information or comments.

In studying this, the following general comments should be considered:

(a) The completed questionnaires, I believe, were generally completed with a good deal of thought and honesty. (One respondent indicated his dissatisfaction with the choices provided. However, the survey form was completed and he apparently recognized the usefulness of such a survey, as he recommended "revision by someone not so devoted to check lists". I have no argument with his viewpoint, and in many cases I have wished that more complete information was available. However, I can think of no simpler form for providing summarized data and the committee intends to provide more objective information in the future).

(b) There were undoubtedly some differences in interpreting the meaning of some of the questions, as exact connotations in communicating is seldom possible.

(c) The planned changes that were checked probably only indicate trends (which is, of course, the most important factor). Undoubtedly many of the companies who replied are considering changes but have not yet made the final decisions, and no information was requested concerning the imminence of these changes or whether they had determined adequate techniques for making the changes or had only decided that they were desirable or necessary.

(d) The replies were from various sources--large companies, small companies, government activities and commercial calibration activities, cloud-nine standards laboratories and working-level calibration groups, newly established groups and ones that have been in business for a long time. Some replies indicated that different systems are in effect in different areas within the organization reported on. Some replies indicated that a very dynamic effort was being made to improve in the areas covered, whereas others have reached a plateau from which they apparently feel only refinements are necessary or practical. However, the concepts involved, if not the techniques, should vary only slightly among these apparently heterogeneous organizations.

(e) The answers indicate that the efforts of the Navy's Calibration Program, as well as those of the other services, have a strong influence on what is being done. Whether this biases the data will have to be decided by each individual. However, whether we completely agree or not, the directives issued by the services have undoubtedly caused us to do things we would not have done or been allowed to do otherwise.

(f) Where possible, I've shown the number of times no answer was given in each area. In most cases, this probably means that respondents believe that this area is not applicable to their operation, that they are unsure what approach to take, or that they haven't yet had time to consider the concepts. However, the relative number of such replies is surprisingly low.

(g) Any bias due to the difference between what is being done and what people believe should be done has been reduced by the anonymity of the replies. However, this factor possibly still causes some bias. Additionally, some bias may result from reporting of conditions that are prevalent in only a part of the organization. However, a number of replies indicated that the companies had additional controls in limited areas but were reporting on their normal systems. As a result, these two factors probably tend to cancel each other.

The organization of the data follows the same outline as the questionnaires and should allow you to pick those areas where you feel your more critical problems exist or where you are interested in how you compare with other organizations. Wherever possible, conclusions are confined to the discussion at the end of each section.

Section II. Recall Systems and Methods

I. Types of Recall Systems Used (A, B & C)

	Present practices		Planned changes to:		Breakdown after changes	
	Number	% of replies	Number	% of replies	Number	% of replies
A. Manual system used	46	63.9	1	-	32	44.4
B. Semiautomatic (punched card, etc.) used	15	20.8	13	-	23	32.0
C. Automatic (magnetic tape, etc.) used	2	2.8	4	-	6	8.3
Manual and semiautomatic	2	2.8	0	-	2	2.8
Semiautomatic and automatic	1	1.4	2	-	3	4.2
Semiautomatic or automatic	-	-	1	-	1	1.4
Manual, semiautomatic, and automatic	1	1.4	1	-	2	2.8
None	2	2.8	-	-	0	0
No answer	3	4.2	-	-	3	4.2
Total	72	-	22	30.6	72	-

The types of manual systems reported were:

Specific types
6 - Kardex
3 - Visirecord
2 - Acme Visible File
2 - McBee Keysort
1 - Unisort
1 - McCaskey
1 - Diebold Tradex
1 - Printed schedule issued and distributed by BUWEPSFLESUPREPCEN

General types
13 - Card file
2 - Revolving card file
1 - Card and chart
1 - Visual inspection record card
1 - Calendar record
1 - Preprinted sheet
1 - "SLIM" dateline 2 weeks recall

Of those replies that indicated the type of semiautomatic or automatic system used, IBM systems were predominant by far. The others specified as being used or planned are:

1 - Remington Rand
1 - IBM card--Remington Rand Sort
1 - GE 225 computer
1 - Philco 2000 at present, IBM 7094 in future

The trend to more automated systems is apparent from the above breakdown. Not one reply indicated plans for less automation than is presently used. Whereas 66.7 percent of the group replying presently use only a manual system or none at all, on implementation of present plans, 51.4 percent of those replying will have a semiautomatic or automatic system in use--almost twice as many as at present.

The answers received did not indicate any leveling off in the degree of automation; the percentage changing from A to B and B to C was substantially the same. However, of the 22 replies indicating a planned change, 20 were from none to A, A to B, or B to C. One company with no current system plans to go to a semiautomatic and one who presently has a manual system indicated plans for an automatic system. Although a gradual transition to more automation is undoubtedly simpler to make, it seems indicated that those who are planning changes (and even those who aren't) should make a careful study of all alternatives before taking the next step. It could well be that, in many cases, what might be termed a double jump would be better.

I am, from experience, sold on the use of an automatic system, not just for recall but for developing the data necessary for many of the controls this committee is covering. At Aerojet we have two substantially identical IBM 1401 magnetic-tape systems. A cost study of one of these systems covering approximately 20,000 instruments indicated that the data processing for recalling instruments and preparing a number of reports and analyses using different sorting criteria costs less than $1500.00 a month including the forms. This is less than the cost of the punched-card system we previously used. We could never begin to justify the number of clerical personnel needed for a manual system that could develop the data we now have, although such data development would probably pay for even this high clerical cost several times over. Fortunately, automated data-processing (ADP) systems are available.

However, even a completely automated system has its drawbacks. Some of the advantages and disadvantages of ADP are as follows:

Advantages of ADP

1. The more automated the system, the greater is the flexibility of reporting, processing, and control that can be built into the system.
2. The data storage and processing capacity increases with the degree of automation.
3. The limitation of 80 digits with most punched cards forces the system to be simplified and organized for optimum results. If this limitation cannot be met, magnetic tape allows virtually unlimited record size.
4. The machine reduces the number of possible clerical errors.
5. Integrated data processing is possible as cross references can be made to other file systems and controls. For instance, the instrument number can be used to check with a property-control system to bring location data up to date.
6. It greatly reduces the clerical help required for a given amount of data or analysis.
7. The cost per unit of information varies with volume, and for larger quantities of information, the cost is relatively low.
8. Its inherent ability to handle large quantities of information results in more centralized controls and consistent application of concepts.
9. Where variations in recording or the specific nature of some information precludes entering it on punched cards or tape, the system can still be programed to isolate those areas where more detailed investigation is warranted.
10. The control obtained is often far more effective and timely than manual systems, as any number of checks can be built into the system. As a result, there is less dependency on intuition.

Disadvantages of ADP

1. The initial expense of programing and subsequent system changes is often high. In addition, the initial recording of great quantities of data to be put on cards or tape may put a large burden on the available manpower.
2. In most cases, going to such a system brings in a data-processing organization that requires a coordinating activity and has its own rules and regulations.
3. You are often competing with other groups for programing and machine time--and all their work is "crash".
4. The simplicity of obtaining reports often leads to reporting of more information than can be put to effective use.
5. Effective use of the machines depends on accurate and precise data recording. The machines still cannot think. Usually, a large number of persons are involved who require instruction, and often the data must be reviewed before the cards are punched.
6. ADP is not always efficient for small operations. However, its applicability depends on the amount of information and/or calculations required rather than solely on the number of instruments.
7. The total cost of data processing is greatly increased if the ADP equipment is only partially utilized.
8. Obtaining special data or instituting changes may require more time and may be more expensive than for a manual system. Programing is usually required as well as coordination with another organization.

We would welcome any additions to this list, as well as other comments, for possible inclusion of such a list in the proposed NCSL "Handbook of Recommended Practices for Standards Laboratories."

II. Notification to user given (D)

Number of days in advance that recall information is given to the user	Present practices		Planned changes to:	
	Number	% of replies	Number	% of replies
30	4	5.6	2	
15 - 30	1	1.4	1	
15	13	18.1	3	
14	2	2.8	-	
10	5	7.0	-	
7 - 10	1	1.4	-	
7	13	18.1	-	
5 - 15	1	1.4	-	
5	3	4.2	-	
3 - 10	1	1.4	-	

Number of days in advance that recall information is given to the user	Present practices		Planned changes to:	
	Number	% of replies	Number	% of replies
3 - 5	0	0	1	
3	1	1.4	-	
2 - 5	1	1.4	-	
2	1	1.4	-	
1 - 2	1	1.4	-	
1	2	2.8	-	
Only when required	1	1.4	-	
No advance notification	12	16.7	-	
No answer	9	12.5	-	
Total	72		7	9.7

Of the seven companies indicating that changes are being considered, only two are currently giving advance notification. Three of the seven indicated that a normal overdue tolerance was allowable and five indicated that longer overdue tolerances were allowable under special situations (30 days in one case). The other two didn't answer the balance of the sheet but indicated they give no advance notice now.

The large variety of answers received and the nature of the few planned changes indicate the following general conclusions:

(a) Most groups believe that advance notification to the user is desirable.

(b) The number of days in advance when notification is given is probably determined more by the system used than by the advance notice required by the user.

(c) As only a few changes are contemplated, there seems to be very little pressure from using groups to give more advance notification. Although there may be some exceptions, it appears likely that only little use is made of this information either because the users don't take the time to use it, can't coordinate all the groups involved, the information is not getting to the right people, or it is not in a usable form.

III. Follow-up Methods (E.1-5)

The complete breakdown of the answers received is as follows:

	Present practices		Planned changes to:
	Number	% of replies	Number
1. By red-tagging overdue equipment	33	45.8	2
2. By use of special expediters/material handlers assigned each department	11	15.3	1
3. By formal notice to user	16	22.2	8
4. By formal notice to department head	23	32.0	4
5. Informal methods only	14	19.4	-
Equipment is removed	1	1.4	
Telephone	1	1.4	
None or not applicable	3	4.2	
No answer	4	5.6	

Of those replies with checks under items 1 to 4, the number indicating 1, 2, 3, or 4 different methods was as follows (presumably all have some informal methods as well):

	Number of replies	% of replies
1 method used	26	36.1
2 methods used	16	22.2
3 methods used	7	9.7
4 methods used	1	1.4
Total	50	69.4

To provide an idea of the order in which these different methods are put into effect, the following breakdown indicates the method used by those that indicated only one method compared with those that use several methods:

Method	Methods used by those checking only one item (26 replies)	Methods used by those checking more than one item (24 replies)
1. By red-tagging overdue equipment	16	17
2. By use of special expediters/material handlers assigned each department	3	8
3. By formal notice to user	2	14
4. By formal notice to department head	5	18

Sixteen questionnaires indicated that changes were planned. Five of these companies use only informal methods at the present time. Two of the five plan to send notices to the department head and three plan to red-tag equipment.

Of the others indicating changes, five apparently plan to stop one of their present methods. All five are planning to add another method as well; one will send a notice to the user instead of red-tagging. Ironically, one plans to stop using expediters in favor of red-tagging and the fifth will stop sending notices to the user and will use special expediters or material handlers.

IV. Normal Overdue Tolerance Allowed (E.6)

Number of days allowed	Present practice	
	Number	% of replies
0	26	36.1
1	8	11.1
2	4	5.6
5	9	12.5
7	6	8.3
10	2	2.8
14	1	1.4
30	2	2.8
Variable	1	1.4
No time limit, but responsibility is shifted to the department head	1	1.4
Individual agreement	1	1.4
Within week scheduled	1	1.4
Unlimited	1	1.4
Doesn't apply	1	1.4
No answer	8	11.1
Total	72	

None of the replies indicated a planned change, although one indicated the company was reviewing its present 7-day allowance.

V. Special Overdue Tolerance Allowed (E.7) (To avoid disruption of test in progress, etc.)

Thirty-seven (51.4 percent of the replies indicated a different overdue tolerance in special cases.
Of these 37 replies, 19 indicated no normal overdue tolerance but some allowance in special cases. The 19 answers in this section are as follows:

Number of days allowed	Number of replies
3	2
7	7
14	2
Varies	1
10% of calibration time (min 2 days, max 7 days)	1
Duration of test	1
Negotiated	1
Special arrangement	1
Depends on individual case	1
Very limited operation	1
By agreement	1
Total	19

The other 18 who replied differently to E.6 and E.7 showed the following variation:

Normal tolerance, days	Special tolerance, days	Number of replies
1	3	2
1	5 (extension by lab supervision)	1
1	7	2
1	10	1
1	As soon as possible	1
2	7	1
2	14	1
2	Unspecified	1
5	10 (plan to reduce to 7)	1
5	14	1
7	14	1
7	30	1
7	Duration of test	1
10	Depends on circumstances--maybe 10-30 days	1
30	30 - 60	1
Individual agreement	7	1
Total		18

Only seven replies indicated that neither a normal nor a special overdue tolerance was allowed.

Section V. Data Collection and Utilization Methods

Besides collecting data for our own purposes, there are a number of DOD specifications, which, to various degrees, require the maintenance of data. Appendix A shows a number of pertinent extracts for those who may not be acquainted with all of them.

I. INSTRUMENT PROCESSING FORMS (Forms used to process instruments once it arrives in the laboratory - A.1-7)

	Present practices		Planned changes to:	
	Number	% of replies	Number	% of replies
1. Part of same form used to recall instrument	10	13.9	4	5.6
2. Special form, separate from recall form, but partially filled out when recall notice is made	12	16.7	0	--
3. Special form initiated by personnel within lab upon receipt of instrument	22	30.6	0	--
4. Form is a card or paper from which punched cards are made	8	11.1	3	4.2
5. Form is a standard manual form, with several copies, which is broken up and sent to various locations	10	13.9	2	2.8
6. Only one form is used for normal instrument processing	38	52.8	0	--
7. ( ) separate forms are used for normal instrument processing	13	18.1	1	1.4
No forms	3	4.2	0	
No answer	6	8.3	--	

Of those companies stating they used more than one form under item 7, the range was 2 to 7 forms and the average was 3.6 forms.

Besides one reply that indicated the company was studying its present system, only seven (or 9.7 percent) showed planned changes.

(a) One company, on the basis of the answers to a series of questions, is just planning a more complete system of processing forms.

(b) Two plan to switch to a form from which punched cards are made.

(c) Three (one of those in b) plan to use part of the same form used to recall the instrument.

(d) One plans to switch from seven forms to one.

(e) One plans to change to a manual form, with several copies, which is broken up and sent to various locations.

As before, the tendency is toward more automated and refined systems. The relatively few replies showing planned changes, compared with the number who are changing their recall system, would support the contention that less attention is paid to the paper work than to a general concept. From an Industrial Engineering viewpoint, all these factors should be studied together with a definite plan detailing what the system is to accomplish. Otherwise, much of the system's effectiveness will be lost and costly system changes will be commonplace.

II. The form/s include data regarding the instrument as follows: (A.8)

		Present practices		Planned changes to:	
		Number	% of replies	Number	% of replies
a.	Manufacturer, nomenclature, model No. and serial no.	64	88.9	2	2.8
b.	Plant accountability no. or ownership No.	54	75	2	2.8
c.	User location	54	75	2	2.8
d.	Reason for submission (normal recall or needing repair due to failure or suspicion thereof)	52	72.2	2	2.8
e.	Date of last servicing	44	61.1	4	5.6
f.	Date of this servicing	63	87.5	3	4.2
g.	Result of initial inspection (visual or otherwise, where signs of damage or mishandling are recorded)	31	43.1	3	4.2
h.	In or out of specification incoming to the laboratory	35	48.6	4	5.6
i.	In or out of specification outgoing from laboratory	34	47.2	4	5.6
j.	Hours to calibrate and make minor adjustments	42	58.3	3	4.2
k.	Hours to repair	38	52.8	1	1.4
	j and k combined	3	4.2	0	--
l.	Parts replaced	47	65.3	0	--
m.	Cost of parts replaced	12	16.7	6	8.3
n.	Number of calibration procedure used	33	45.8	2	2.8
o.	Name or number of calibration technician	52	72.2	1	1.4
p.	Temperature and humidity condition	25	34.7	5	7.0
q.	Calibration equipment used to perform calibration	32	44.4	2	2.8
* r.	Actual measurement data as recorded during calibration process	41	57.0	2	2.8
	None	4	5.6	--	--
	No answer	1	1.4	--	--

* A number of the replies indicated that 'r' applied only part of the time, such as "working standards only".

Seventeen replies indicated some change was expected. Apparently no one planned to record less data than at present (which should please C. Northcote Parkinson).

Two of these changes were planned by groups not now recording any of the listed data. Seven of these respondents planned to add only one item. These included five of the six replies that indicated plans to record the cost of parts replaced. As these five recorded most of the other data shown, this is undoubtedly believed to be a refinement that takes a while to put to practical use and then only by those that have gone deepest into this type of control. Comparing item m above with the answers in items D.1 and D.2 shows a close correlation. Although the numbers are relatively small, the trend is again apparent. It might be well for those not now maintaining replacement parts data to study this area more thoroughly.

III. Measurement-Data Recording (B)

	Present practices		Planned changes to:		Breakdown after changes	
	Number	% of replies	Number	% of total	Number	% of replies
1. Is not done. The technician is expected to either follow a procedure or use his best judgement	16	22.2			10	13.9
2. Actual data is not recorded, but a preprinted check list is used for each test point and allowable tolerance is indicated	13	18.1	3	4.2	13	18.1
3. Actual data is not recorded, but the technician indicates how far in or out-of-tolerance each reading is by means of a "tolerance range" table and coding system	3	4.2			2	2.8
4. Actual data is recorded used preprinted forms showing each test point and allowable tolerance	34	47.2	7	9.7	41	57.0
5. Actual data is recorded without use of special forms, and reliance is placed upon technician to determine test points and tolerances	20	27.8			20	27.8

A number of the replies indicated more than one of the listed methods of recording data. Nevertheless, the answers to this section surprised me, even after making allowances for some "over-reporting" of what is being done. When all of the indicated changes are put into effect, over two thirds of the reporting organizations will be recording the actual calibration data. Only 13.9 percent indicated that item 1 applied, and several of these had checks for other items as well (however, some allowance should also be made for the few who didn't answer this section, as they probably tend to fall under item 1).

All of the nine replies indicating planned changes are "upgrading" their systems by recording more data. Again, nobody, apparently, is planning to record less data in the future.

There are several arguments in favor of formal recording of the data in some form.

(a) It ensures that all required points are checked.
(b) It provides the supervision with feedback on how well the technicians understand the applicable procedures and tolerances.
(c) The technician, if he's doing his job, is probably writing down much of the data on a piece of scrap paper anyway.

However, formal recording of data introduces more paperwork into the system, and analysis of the data is virtually infeasible. Many of the people I've talked to have experienced an increase in workload of 25 to 50 percent when the technicians were required to record data. Although these technicians are undoubtedly doing a more thorough job, they also probably hesitate to use their own discretion even when it is justified.

Some of the comments received in connection with the above discussion are as follows:

(a) The volume of instruments is such that the supervisor can review all data taken to assure the quality of the calibration.
(b) Historical records are maintained of all equipment calibrated.
(c) Actual data is not recorded, but calibration procedures are used that indicate tolerance limits. On special need or request, data is taken and supplied to users.
(d) Actual data is recorded where accuracy is 1 percent or better--except on secondary or working standards (several similar comments).
(e) The check list is part of the service manual.
(f) Adequate data is the only proof of a completed task. Reputation alone is not always sufficient.
(g) Detailed written procedure may specify recording check points or other information on IBM card for critical production test equipment. Special data sheets specified for standards.
(h) Procedures are followed. Only out-of-tolerance information is recorded.
(i) Working on method of recording pertinent data. Data on mechanical equipment is recorded.
(j) Paragraphs 2 and 4 are used depending on common sense requirements.
(k) In effect, all apply. The recording procedure is determined by the nature of the test instrument or working standard.

IV. User and cognizant inspection personnel are regularly notified when an instrument is found out of specification incoming to the laboratory by: (C)

	Present practices		Planned changes to:		Breakdown after changes	
	Number	% of replies	Number	% of replies	Number	% of replies
1. Informal discussion	32	44.4	0	--	21	29.2
2. Formal notification system	16	22.2	16	--	32	44.4
3. Not at all	13	18.1	0	--	8	11.1
Both 1 and 2 apply	4	5.6	0	--	4	5.6
Both 1 and 3 apply	3	4.2	0	--	3	4.2
Both 2 and 3 apply	1	1.4	0	--	1	1.4
"Not a problem"	1	1.4	--	--	1	1.4
No answer	2	2.8	--	--	2	2.8
Total	72		16	22.2	72	

Obviously, the general concensus of opinion is that the user should be told whenever instruments are found to be discrepant. The limiting factors are probably the development of a simple method for notifying the user and for assuring that corrective action, in some form, is taken. Nevertheless, nearly 50 percent have decided that they are in a position to realize benefits from this type of communication.

Eleven groups who now notify informally and five who provide no notification plan to institute a formal notification system.

However, six of the eight remaining who do not provide feedback specifically indicated they did not plan to make a change. Two of the three indicating that both 1 and 3 applied also indicated that no changes were planned. Such definiteness was not apparent in most of the categories of the survey.

V. Data Utilization Includes: (D)

D.1 Adjustment of instrument calibration intervals based on summaries of data

a, b, & c

	Present practices		Planned changes to:		Breakdown after changes	
	Number	% of total	Number	% of total	Number	% of total
a. Reviewed annually*	14	19.4	4	5.6	16	22.2
b. Reviewed semiannually*	9	12.5	6	8.3	16	22.2
c. Reviewed only when a particular instrument is "suspect"	28	38.9	0		22	30.6
Reviewed quarterly	2	2.8	0		2	2.8
Constant review	2	2.8	0		2	2.8
Following each calibration	1	1.4	0		1	1.4
None	8	11.1	0		5	7.0
Comments only	3	4.2	--		3	4.2
No answer	5	7.0	--		5	7.0
Total	72		10	13.9	72	

*Most of those checking a and b indicated they also reviewed when a particular instrument is "suspect".

d & e

Thirteen replies indicated that they also adjusted intervals based on the concept of increasing intervals if less than a specified percent were found discrepant and on the concept of decreasing intervals when the proportion of out-of-tolerance instruments is beyond a specified limit.

Seven additional replies indicated they would institute this concept in the future.

Of these, 13 replies stated the action limits that were used:

% for lengthening intervals	% for decreasing intervals
2	5
5	5
5	5
5	15
5	20
5	-
10	10
10	20
10	30
10	40
15	40

(The following two are also included although there is apparently a typographical error)

90	20
80	20

One additional reply stated that action limits depended on "level of calibration and how far out".

D.1.f - Comments

The following pertinent comments were submitted pertaining to calibration intervals:

a. Instrument history is reviewed following each calibration to determine the most desirable interval.
b. Most cycle rates are determined for us by the Navy "SLIM" manual. We are forced to comply.
c. Intervals are primarily fixed, but semiannual reviews are made anyway. Interval adjustments are made in a few cases dependent on instrument and application.
d. Adjustments are made by Navy program management.
e. Review of "Cal" results are surveyed on a continuing basis, with emphasis on types showing frequent out-of-calibration findings.
f. Instruments are reviewed per make and model, and calibration intervals are set according to type, quality, and usage.
g. Lab in organizational stage, cannot comment (several such replies).
h. This depends on device and use and is usually a decision arrived at by phone discussions.
i. Long-range plans call for establishing calibration intervals based on actual use and not time.
j. Extension of interval is negotiated.
k. Individual instrument histories are reviewed at the time of each calibration. Adjustment of calibration interval is somewhat intuitive and is not usually categorical.
l. When contractual calibration periods are shorter than our company-determined periods, the contract period takes the place of the company periods; otherwise company periods are used.
m. This review is made by the Q. C. Engineering evaluation group and changes are coordinated with Air Force representatives before authority is granted for the change. (Editor's Note: I hope this doesn't catch on.)
n. IBM process tabulates all instruments and their repair code semiannually. This information is used as basis for "what to investigate".
o. Normal calibration period of 90 calendar days, no period exceeds 6 months.
p. Adhering to calibration periods established under the BUWEPS Standards Program. We have shortened some periods but will not lengthen them.
q. No regular basis.
r. The above percentages are guides. Further analysis is based on such items as use of instruments, type of discrepancies, possible causes, and other solutions.
s. Submitted to Navy, BUWEPS, Pomona, California, for evaluation and subsequent rescheduling instruction.
t. Adequacy of interval period to be determined by new system.
u. Intervals are based on instruments usage and accuracy necessary.

No factor has a greater influence on a laboratory's workload than calibration intervals, and except, possibly, for the technical competency of the calibrations, none has a greater effect on measurement reliability. Ideally, these two opposing factors must be delicately balanced to arrive at an optimum recycling time. A great deal can be lost by gambling on intuitive judgement, and the gains possible from large expenditures of manpower and technical competency are tremendous.

The replies on the questionnaire indicate that the standards and calibration activities throughout the country have taken a pretty inconsistent, and even sloppy, approach in the establishment of calibration intervals. The most common analytical approach is to adjust these intervals on the basis of the proportion that are found discrepant. However, only a relatively few have specified even preliminary criteria for this decision-making process. And these criteria are probably the most inconsistent in the survey.

According to the replies in section V.A.8.h, less than half the organizations replying are even recording data as to whether or not instruments are in tolerance when received.

This conclusion is supported by an almost complete absence of papers published on the subject. However, the considerable interest being expressed for others to present papers on this subject at least indicates there is some awareness of the inadequacies in this area.

There is no intent to oversimplify the problem. In fact, it may well be the most complex one we are faced with. The solution depends on the adequacy of virtually all the other controls we have established--adequacy of calibration procedures, accuracy ratios, correct data reporting and analysis, promptness of recall, etc. It depends similarly on factors not under the control of the calibrating activity--what accuracy is required in use, whether specifications are realistic, whether the user questions the performance when he should, whether there are built-in checks to guard against erroneous measurements, whether the instrument is subjected to heavy use or used only sporadically, etc.

Any solution must be a compromise and assumptions must be made, at least in our present state of proficiency. And the complexity of even partial solutions requires a concerted effort by many people, to the extent that it is probably beyond the present capabilities of NCSL or other technical societies to provide realistic criteria.

With this in mind, I propose that NCSL officially recommend that DOD investigate the use of statistical tables for establishing and controlling calibration intervals. These could be similar to Mil-STD 105A and similar sampling tables. In fact, the problems outlined above are virtually the same as those encountered by inspection activities in applying sampling techniques--and such sampling has become one of the most widespread statistical applications in industry.

Two examples of tables that might be derived, one comparable to single sampling and one comparable to sequential sampling, are:

(a) To establish tables that would provide upper and lower action limits for extending or shortening intervals based on the acceptable percentage of discrepant instruments, the actual percentage of discrepant instruments, and the number of calibrations on which the data is based.

(b) To determine the practicality of and to develop tables, based on the acceptable percentage of discrepant instruments, that would provide a maximum calibration interval for items found within specifications and a reduced interval for instruments found to be discrepant. An example of this would be an instrument that has been calibrated several times and found to meet specifications. It would be recalled every 6 months. However, if discrepancies were found, it would be recalled in 2 months. If, after two or three recalibrations, no further discrepancies were found, the interval would again be extended to 6 months until it was again found discrepant. This would materially assist in calibrating those items that, through use or misuse, were often discrepant while not overcalibrating like items used with care or infrequently.

VI. Summaries are made to determine (D.2)

	Present practices		Breakdown after changes	
	Number	% of replies	Number	% of replies
Number of companies making listed summaries	40	55.6	51	70.8
On special request	1	1.4	1	1.4
None	21	29.2	10	13.9
No answer	10	13.9	10	13.9
Total	72		72	

		Present practices		Planned changes to:		Breakdown after changes	
		Number	% of replies	Number	% of replies	Number	% of replies
a.	Average calibration time per instrument and instrument category.	35	48.6	9	12.5	44	61.1
b.	Average repair time per instrument and instrument category.	23	32.0	11	15.3	34	47.2
c.	Replacement parts required per instrument.	11	15.3	9	12.5	20	27.8
d.	Parts cost per instrument	9	12.5	11	15.3	20	27.8

Of those companies making the listed summaries, the following breakdown indicates the number making 1, 2, 3, and 4 of the listed summaries.

	Present practices		Breakdown after changes	
	Number	% of replies	Number	% of replies
1 of listed summaries	18	25.0	17	23.6
2 of listed summaries	8	11.1	11	15.3
3 of listed summaries	8	11.1	9	12.5
4 of listed summaries	5	7.0	13	18.1
Total	39	54.2	50	69.4

(One reply indicated they were making summaries but didn't break them down.)

The following breakdown indicates the types of summaries made by those reporting only one summary:

	Summaries made by those making or planning only one summary
a. Average calibration time per instrument and instrument category.	15
b. Average repair time per instrument and instrument category.	2
c. Replacement parts required per instrument.	2
d. Parts cost per instrument.	1

D.2.e Comments:

Pertinent comments given that are applicable to this section include:

*(a) At the time of purchase of a new instrument, a survey is made of the instrument and parts list. Critical parts are then ordered and stocked. The spare-parts stock is checked regularly and parts reordered according to a min-max label on parts bin. (No checks were shown on this reply.)
*(b) Adjustments made by Navy program management. (No checks were shown.)
(c) Also to determine manpower requirements and repair parts stock.
*(d) No summaries--too early. (No checks were given or assumed--several similar replies.)
(e) Have the information for a, b, c, and d but have not been able to summarize. (Checked under planned.)
(f) Quarterly history reports are made by model number showing percentage out of tolerance by reason for service and average hours to service, plus detailed listing of instrument by serial number as requested. Quarterly reports are also issued showing breakdown by technician.
(g) Summaries are made to reflect instrument calibration and maintenance costs, reliability, and limit of useful life. (No checks were shown.)
(h) Costing of items 2d above and 3a below is under consideration, but method of separating from other test laboratory costs has not been determined.
*(i) In our small laboratory, lack of clerical and technical personnel does not allow for making summaries, studies, and elaborate analyses.

*These comments are also applicable to D.3.

VII. Special studies are made to determine: (D.3)

	Present practices		Breakdown after changes	
	Number	% of replies	Number	% of replies
Number of companies making listed special studies	35	48.6	41	57.0
None	22	30.6	16	22.2
No answer	15	20.8	15	20.8
Total	72		72	

	Present practices		Planned changes to:		Breakdown after changes	
	Number	% of replies	Number	% of replies	Number	% of replies
a. Annual owning cost per instrument.	7	9.7	5	7.0	12	16.7
b. Retirement time for instruments.	14	19.4	9	12.5	23	32.0
c. Stock level and kind of replacement parts.	23	32.0	6	8.3	29	40.3
d. Need for derating of instruments due to constant out-of-tolerance condition.	18	25.0	9	12.5	27	37.5

Of those companies making the listed summaries, the following breakdown indicates the number making 1, 2, 3, or 4 of the listed special studies.

	Present practices		Breakdown after changes	
	Number	% of replies	Number	% of replies
1 of listed reports	15	20.8	14	19.4
2 of listed reports	12	16.7	10	13.9
3 of listed reports	5	7.0	7	9.7
4 of listed reports	2	2.8	9	12.5

The following breakdown indicates the types of special studies made by those reporting only one such study (either currently or planned.)

	Special reports made by those making or planning only one change
a. Annual owning cost per instrument	1
b. Retirement time for instruments	2
c. Stock level and kind of replacement parts	6
d. Need for derating of instruments due to constant out-of-tolerance condition.	8

Some of the comments and other special studies shown in this area are as follows (some of the comments shown under D.2 apply here as well):

(a) Only to determine applicability. (Item d was the only one checked.)
(b) If cost of a repair exceeds 60 percent of replacement cost, repair is not done.
(c) Special measurement area studies, such as phase angle studies, microwave power studies, VSWR, and attenuation studies. (Items c and d were checked.)
(d) Studies are made only when and if need becomes evident.
(e) Item A is studied strictly on an informal basis; item d is not applicable, as we either repair or scrap any instrument we cannot make meet manufacturer's specs.
(f) Out-of-tolerance instruments are not used. (Several similar replies.)
(g) Need for personnel training, need for different type of equipment, change of quantity in an area.
(h) Repair service versus instrument cost.
(i) Item c and d are under constant appraisal by lead engineers of the laboratory sections.
(j) Retirement based on obsolescence, wear, or availability of improved equipment. (No checks were made.)
(k) One reply indicated studies for utilization, distribution of costs, and shop efficiency. (No checks made in items a through d.)

Conclusions

The purpose of the survey and of this report was to gather information on what the various standards and calibrating groups throughout the country are doing. The value of the data and the importance of the conclusions will depend primarily on the needs and interests of each organization involved.

Nevertheless, several general conclusions can be drawn from this data:

(a) The desirability for controls in the areas covered is generally accepted. The trend is definitely toward more controls with greater refinements. This is a natural culmination to the emphasis placed on calibration in recent years and to the increased activity of technical societies in disseminating information on the concepts and techniques used for such controls.

(b) Although there are necessary differences between organizations, the concepts and, to some extent, the best techniques vary far less than our products or who we work for.

(c) The necessity for providing automated data-collection methods is virtually essential in those operations of any significant size. This need does not, as is commonly stated, depend solely on the number of instruments to be controlled. Rather, it is a function of:

  Number of instruments x number of calibrations per instrument x the quantity of information required per calibration x the number of computations to be made per unit of information.

(d) Considering its importance to our operations and the lack of consistent and accepted approaches, the greatest weakness in our controls is in the establishment of calibration intervals. To overcome this weakness, a great deal of effort must be expended to develop realistic concepts and techniques.

Appendix A. Selected DOD Requirements For Maintaining Calibration Data

1. "Standards Laboratory Information Manual", BuWepsRep, Pomona, Calif.
   a. Par. 2, Page 5.9-4--"A calibration report form is to be utilized by each laboratory and will be completed for each calibration performed."
   b. Par. 3, Page 5.10-1--"When utilizing all ICP's, the Calibration Report should be filled out exactly in accordance with the Calibration Report contained in the last page of each ICP. This will guarantee consistency of data recording, consistency of calibration techniques, and simplified analysis of calibration data."

2. MIL-Q-21549A (NOrd) "Quality Assurance Program Requirements for Fleet Ballistic Missile Weapon System Contractors"
   a. Par. 3.5.5.3 Maintenance of Inspection and Test Equipment--"Test or inspection equipment shall not be used without written notification to the Special Projects Office, or its designated representative, if the prescribed date for recalibration has been exceeded or the equipment is found to be outside of established accuracy limits. Records shall be maintained on the recalibration status of all measurement equipment. Records shall include adequate identification of each specific measurement equipment, an indication of the condition of the equipment and corrections made at each check or recalibration. There shall be an indication on each specific measurement equipment, or readily available at the location of the equipment, indicating the date of last check and date next calibration is due. Variables data shall be maintained and an analysis of this data shall be performed to indicate trends of wear, deterioration and maintenance in order that maintenance and recalibration procedures and schedules may be realistically revised to assure required accuracies of the equipment."

3. MIL-Q-9858 "Quality Control System Requirements"
   a. Par. 3.16 Quality control records--"The contractor shall maintain adequate records throughout all stages of contract performance of inspections and tests, including checks made to assure accuracy of inspection and testing equipment and other control media."

4. MIL-C-45662 (Ord) "Calibration of Standards"
   a. Par. 3.1.7 Application and Records--"The application of the above requirements will be supported by records designed to assure that established schedules and procedures are followed to maintain the accuracy of all standards."

5. U. S. Air Force Specification Bulletin Number 520, "Calibration and Certification of Measuring and Testing Equipment"
   a. Par. 7 Scheduling and Record Keeping--"Schedules shall be established and records maintained to assure that calibration of measuring and testing equipment is performed periodically to assure continued accuracy."

6. NASA Quality Publication 200-2 "Quality Assurance Provisions for Space Systems Contractor"
   9.7 Records--Records shall be maintained on the recalibration status, condition, and corrections or repairs for each inspection, measuring, and test equipment. Variables data shall be maintained and an analysis performed to determine trends of wear, deterioration, and adequacy of maintenance. Procedures and schedules shall be realistically revised accordingly.

7. BuWepsInst. 4355.5--"Bureau of Naval Weapons Calibration Program"
   a. Par. 6.c.2.d--The maintenance of records which will show the calibration status of all measuring equipments and standards and will insure positive recalibration at specified intervals of all test equipment utilized in support of the prime contract.

8. MIL-C-45662A--"Calibration System Requirements"
   a. Par. 3.2.6 Application and records--"The application of the above requirements will be supported by records designed to assure that established schedules and procedures are followed to maintain the accuracy of all measuring and test equipment, and supporting standards. The records shall include a suitably identified individual record of calibration or other means of control for each item of measuring and test equipment and measurement standards, providing calibration interval and date of certification of results of last calibration. In addition, the individual record of any item whose accuracy must be reported via a calibration report or certificate will quote the report or certificate number for ready reference. These records shall be available for review by authorized Government personnel."

9. MIL-R-25534A "Rocket Motors, Aeronautical, Qualification Test for.
   a. Par. 4.1.1.1 Accuracy of Data.--"---------All apparatus shall be calibrated often enough to insure that this degree of accuracy is maintained. Calibration records shall be retained by the testing agency for 2 years, and furnished to the procuring activity upon request.

NBS Miscellaneous Publication 248
Proceedings of the 1962 Standards Laboratory Conference

Session 7. Calibration Recycle Analysis and Work Load Control

Paper 7.5. Data Processing for Control of Instrument Maintenance and Calibration

J. W. Rodgers, Jr.*

This paper will discuss the total utilization of data collection throughout all aspects of standards laboratory management and operation as experienced and implemented by Bendix Radio. The systems employed, the reasons for their selection, and the experience of seven years of operation under such a system will be presented as an example of the application of those aspects discussed in previous papers of this session. The use of these systems has resulted in minimizing clerical staff and provides the means to evaluate overall performance of laboratory operations in regard to the cost of maintenance and calibration of test equipment. These evaluations result in establishing cycles of recalibration, methods of employing manpower more effectively, evaluations of savings for justification of new test equipment purchases, and many other by-products. Without effective use of data processing, much of this information has been found to be either too costly to obtain or simply unavailable. The paper will emphasize when data processing should be used, the development of a punched card program, the establishment of calibration cycles based on instrument rejection rate, the evaluation of unit cost or standard hours of maintenance and calibration and feedback to determine manpower requirements and overall work load control.

1. Introduction

In this morning's session, you have heard methods of operations for a standards laboratory including utilization of manpower, planning of workload, how often to establish a recalibration program, limited and restricted calibration, and other aspects of standards laboratory controls.

It is my purpose to present a data process system that has been in use at the Bendix Radio Division for the past seven years. This system is supplying such vital information as standard hours per instrument for repair and calibration for determining workload and manpower needs; acceptable levels of recalibration, used to adjust recalibration scheduling on the basis of these levels; evaluation of equipment failures to determine equipment manufacturers' acceptability; employee efficiency; and other by-products.

The advantages and disadvantages of this system have been carefully weighed to minimize the amount of recording of clerical work by the individual calibrator and repairman and still obtain the more vital information as a permanent record. I have no doubt that a more comprehensive system can be developed, but one must weigh the cost of such a system in man hours required and the gains derived.

In order to intelligently approach this problem, the first question that we must ask ourselves is, "Is data processing equipment available within our plant or would such work have to be contracted?". Cost of such work under each condition can have quite a variance. As an example, when using the system presented, the crossover point for machine tabulation versus hand tabulation when an in-plant facility is available is estimated at 165 units of repair and calibration work at a cost of approximately $35.00 direct labor. Cost of this same work being processed by a business machine service would establish the crossover point at a volume of 700 units and a cost of $145.00.

One must first determine what his needs and requirements are and then establish the method based on facilities available and the current clerical workload and cost.

In the case of the Bendix Radio system, we currently process approximately 2000 instruments per month with an in-plant data processing facility. The cost of our data processing is $110.80 per month.

---

*Bendix Corporation, Bendix Radio Division, Baltimore, Md.

A similar clerical function would cost approximately $210.00. The clerical function obviously would be quite limited in regard to special reporting. The cost of such reporting would, in most cases, be prohibitive due to excessive cost and poor timing. When dealing with hand tabulation of special reporting the cost would generally be approximately $70.00 per report. Special reports would cost approximately $20.00 per report using data processing technique.

2. Bendix Radio System

Exhibit "A" illustrates the work card currently in use and the I.B.M. 5081 punch card used by the Standards Laboratory to record instrument history. Coding has been used to maintain all information on an 80 column card. Whenever coding is used these codes have been devised, where possible, to resemble commonly used abbreviations, since such abbreviations and standardization are more easily recognized and remembered by calibration and repair personnel.

Exhibit "A" also indicates the spacing allowed on the punch card for each item numbered on work ticket.

Definition of instrument type can be (found in exhibit "B"). Farther expansion of this coding could be accomplished by use of alphabetical coding as opposed to numerical if such expansion is deemed necessary.

The manufacturers code is obtained by the use of "SLIM", U.S. Naval Inspector of Ordnance Publication, appendix F. Code standardizing is always advantageous when possible.

Items 3 through 9 are self explanatory.

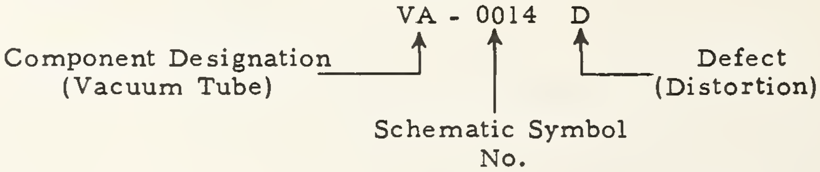
Item 10 indicates the disposition code and is coded as presented:

Reject and Acceptance Code

The following reject and acceptance coding is used by Bendix Radio Division personnel in processing work.

Digit No.	Disposition	Definition
2	Reject	Recalibration *not in excess of one month* of due date (calibration only)
4	Reject	Equipment rejected from repair personnel of Department 953
5	Reject	Recalibration *in excess of one month* due date (calibration only)
6	Accepted	Recalibrated with *minor adjustment* necessary (calibration only)
7	Accepted	No adjustment necessary (calibration only)
8	Accepted	*Minor adjustment* necessary *from repair* section
9	Accepted	*No adjustment* necessary *from repair*

The fault code is developed as outlined in this exhibit:



You will note that this coding has been derived to use standard schematic component designation numbers used by the manufacturers to obtain maximum direct referencing without the use of decoding sheets.

It should be noted that by the use of this system, the primary fault of each instrument is recorded for future evaluation. The primary fault being the predominant defective component that caused the instrument failure.

The complete fault coding is attached to reprints.

Now that we have established the method and system to be used, i.e., DP or clerical, and the type of information to be recorded, one must ask himself what information is necessary to obtain from these records for the operation of a standards laboratory.

The Bendix Radio Standards Laboratory has a total of nine reports developed from this data, four of these reports on a monthly basis, three on a semi-annual or annual basis and the remaining two as required.

I will discuss these reports in order of their importance to our standards laboratory operation, the first being the calibration recycle report (ref. exhibit "B"). This report is most important since it establishes cycle of calibration and thus determines workload. As noted from this exhibit, there are three evaluations on rejects; rejection of equipment that are not over one month expiration, rejections over one month of expiration, and rejection from our repair shop. It has been found necessary to make this separation to more properly evaluate a true calibration cycle. Without such separation, a true rejection ratio cannot be established. We obviously cannot include such factors as equipment rejected from the repair facilities and rejects of equipment that has been subjected to excessive use without recalibration.

On the basis of past failure rates, we have established an acceptable range between 80 percent and 90 percent. Continued acceptability levels below the 80 percent region is reason to increase our calibration cycle. Continued acceptability above 90 percent level is reason to decrease the calibration cycle.

The level of acceptability becomes a problem of economics as well as reliability. I am sure that most of you would agree that 99.9 percent would be more in line with what we would like to achieve, but to do so would result in greater downtime than use time and excessive calibration cost.

This report is issued on a semi-annual basis, time being revised as the sampling quantity indicates. Small samples would obviously lead to false conclusions and therefore are discounted until adequate sampling can be achieved.

A revision of calibration scheduling at Bendix Radio on August 30, 1961, resulted in a lengthening of the calibration cycle resulting in a 2105 hour reduction of calibration time and an annual savings to the Division of $6315.00. There has been no reluctance on the part of our customers to extend this calibration cycle with the facts presented by use of the data processing method.

3. Vendor Rating

By a slight modification of the calibration recycle report, data is obtained to evaluate vendor reliability. To obtain this information, a sub-division is made of instrument type to include manufacturers. Such a report would include instrument type as presented (in exhibit "B") and sub-divided by the three-letter code of manufacturers. Totals of repair, calibration and rejection would be presented as previously shown.

No attempt has been made to evaluate manufacturers by specific model number. However, this can be done by further sorting the data cards to include this information. It is again important to point out that small samples can be misleading. Therefore, sampling should be taken over a time duration to obtain reasonable quantities. Usually one year or over is a reasonable sample when dealing with vendor evaluations.

Example:

Instrument type		Manufacturer	% Reject 1 Mo. $\frac{2}{2+6+7}$
1	Amplifiers & receivers	GEB	8.0
		TEA	5.0
			6.0 Total
5	Attenuators	BOA	6.0
		DAC	7.2
		GEB	5.0
		HEA	9.5
			7.8 Total
10	Audio oscillators	GEB	10.0
		HEA	9.8
			9.9 Total

This information is presented as examples and in no way infers the actual quality of manufacturers presented. Their identification is to present only an example of the methods used to obtain this data.

Labor Distribution

The next most significant report is the labor distribution report.
Example:

Department	Units of calibration	Total hours of calibration	Hours of special calibration	Avg. calibration time	Total hours of repair	Subtotal of repair and calibration hours	Hours of special repair	Subtotal of special repair & calibration hrs	Units of repair	Units of repair & calibration	Avg. repair time
151-0	6	10.3	0	0.94	23.5	33.8	10.0	10.0	3	5	2.9
161-0	53	12.6	0	.2	58.9	71.5	0	0	2	7	6.5
211-0	6	77.1	0	1.68	191.8	268.9	0	0	22	40	3.1
251-0	10	11.3	12.0	0.95	8.0	19.3	0	12.0	0	2	4.0
Totals	75	111.3	12.0	.86	282.2	393.5	10.0	22.0	27	54	3.5
	↑	↑		↑	↑				↑	↑	↑
										↑	

The most valuable information received from the standpoint of standards laboratory operations is the standard hours of calibration and repair on a total workload basis. This is a yardstick in determining overall efficiency within a repair and calibration activity and is developed on a monthly basis. Action can be taken to improve work performance as needed.

In the case of Bendix Radio Division, this report is also used to distribute operating cost of the standards laboratory to the product groups using the service, hence the sorting by department number.

4. Labor and Workload Forecasting

The labor and workload forecast report (exhibit "C") is used for evaluation of repair and calibration time. Such a report indicates cost per unit for repair and calibration and manpower requirement to maintain this equipment.

One must be aware that active equipment use within a facility may bear no resemblance to present work capabilities within a standards laboratory. Therefore, this information generally cannot come from past history of standards laboratory workload. This information, in the case of Bendix Radio, was obtained from the Instrument Record Section. Such records indicate actual instrumentation assignments to projects within the Division and therefore represent actual forecasted workload for the standards laboratory. New equipment input, obsolete equipment dispositions, and instrument storage would adjust such active equipment requirements. Such records are maintained by the Instrument Record Section and supplied as requested.

Budgetary requirements can be developed from such information to facilitate determination of manpower requirements for the standards laboratory. It is well to note that this manpower requirement now becomes factual as opposed to educated guessing and is extremely valuable in convincing higher management of existing and forecasted needs.

5. Calibration Cycle and Recall

From the information given in the labor and workload forecasting report, total workload can be forecast. However, spreading of this workload equally on a monthly basis becomes another matter. This can be accomplished by the use of the calibration cycle and recall report.

This report as presented has been separated by department and is published each month to notify using departments of calibration expiration. By the use of this report, workload forecast can be projected for six months, one year, or beyond when necessary.

The calibration due code is essentially an alphabetical code by month and is presented in the reprint. This code has been arranged so that the month code and recalibration code duplicates each 22 years. This system was selected in place of day, month, year to minimize card space.

Example:

Instr. code	Mfgr. code	Model	Asset No. Tool No. etc.	Date last calibrated	Dept.	Calibration due code
10	HEA	200CD	149-7640	704	455-0	C
10	BEA	52604	149-6705	631	476-2	B
20	GEB	1409K	476957101	931	476-2	C
5	DAC	RFB55050	149-2076	320	462-2	I
45	WEC	433	149-3078	520	462-2	F
60	TEA	545A	149-4070	620	462-2	B
70	BOA	202	149-2077	728	462-2	C

In order to forecast peaks and valleys of workload, one need only to develop a workload forecast. Such a workload forecast is data programmed by use of summary cards indicating time factor by instrument. Example:

July

Instr. code	Quantity	Time factor per unit	Time per instr. type	Calibration due code
1	51	3.91	199.41	A
10	30	2.601	78.030	A
20	21	1.606	33.726	A
30	18	3.005	54.09	A
60	25	3.162	79.05	A
70	30	4.098	123.84	A
	175 units		568.146 hr	

$$\begin{array}{l} \text{Total workload.} \\ \text{July work hr per man} \end{array} \quad \frac{568.15}{157.5} = 3.6 \text{ men.}$$

August

Instr. code	Quantity	Time factor per unit	Time per instr. type	Calibration due code
5	21	1.629	34.209	B
10	8	2.601	20.808	B
41	38	1.349	51.262	B
43	40	0.530	21.2	B
46	38	0.983	37.354	B
60	10	3.162	31.62	B
75	5	4.098	20.49	B
	160 units		216.943 hr	

$$\begin{array}{l} \text{Total workload.} \\ \text{Aug. work hr per man} \end{array} \quad \frac{216.9}{172.5} = 1.26 \text{ men.}$$

As can be seen from the example presented, the quantity of work has no resemblance to man-hour requirements for a given month and thus the need for evaluation.

The six reports just covered are basic tools for management for determining workload, cost, overall efficiency, and quality of instrument calibration and repair.

6. Detail--Efficiency and History

In order to improve the reliability, efficiency, and work methods, it may become necessary to have more detailed information on a laboratory operation. This is obtained at Bendix Radio by a data programmed detail evaluation of employee and instrument problems. By the use of such information, management can minimize the required direct supervision. To do so merely requires the evaluation of the data obtained.

There are three such reports obtained for this purpose; efficiency reports of calibration personnel and repair personnel, and the evaluation of instrument failures, including replaced components.

7. Calibration Personnel Efficiency

Example:

Employee class. plant and shift	Employees clock number	Reject #2 in date calibration	Reject #4 from repair	Reject #5 overdue calibration	Total reject (1) #2 + #4 + #5	Acceptance #6 minor adj. calibration only	Acceptance #7 No adj. calibration only	Acceptance #8 minor adj. from repair	Acceptance #9 No adj. from repair	Total all work processed (2)	% of rejects to total processed = $\frac{(1)}{(2)}$
A	1034	1	1	2	4	10	60	3	10	87	4.6%
A	1218	0	1	1	2	9	68	8	6	93	2.1%
B	1209	2	4	3	9	14	125	13	20	181	5.0%
B	3913	1	2	6	9	18	102	7	15	151	6.0%
D	4528	1	2	1	4	8	53	4	11	80	5.0%
S	1219	2	0	0	2	2	20	0	0	24	8.3%
U	1247	1	0	2	3	12	59	2	13	89	3.4%
Total		8	10	15	33	73	487	37	75	705	4.7%

The calibration efficiency report has been extremely valuable in determining calibrator efficiency in terms of monthly work performance of the individual calibrators.

To properly evaluate performance, it becomes necessary to evaluate on the basis of similar work, since it is obvious that there will be less work performed in terms of quantity when a calibrator is working on high frequency signal generators, counters, etc., as opposed to a calibrator performing 2 percent meter calibration.,

Percentage rejection by calibration and repair personnel in many cases present similar patterns.

8. Repair Personnel Efficiency

A similar method of recording quantity and quality of work is employed for the repair personnel.
Example:

Employees class. plant and shift	Employees clock number	Acceptance minor adj. #8	Acceptance No adj. #9	Total acceptance #8 and #9	Total all units calibrated (2)	Reject #4 (1)	% of total acceptance = $\frac{(1)}{(2)}$	Repair No calibration necessary	Total of all jobs produced
A	1212	1	19	20	21	1	4.8%	10	31
A	1261	0	29	29	29	0	0 %	2	31
B	1274	2	22	24	28	4	14.3%	12	40
D	1271	1	25	26	26	0	0 %	0	26
E	1233	0	35	35	36	1	2.8%	18	54
E	4639	0	40	40	42	2	4.8%	15	57
Total		4	170	174	182	8	4.4%	57	239

These detailed reports are used in discussing efficiency with individual employees by their direct supervisor.

9. Instrument History Report

The final routine report is known as the instrument history report. This report essentially covers all items programmed on the punch card as previously described in exhibit "A" and is developed annually.

It might be well to note that this report has been arranged in such a manner as to program instrument repair and calibration records of like serial number and record single space. Double spacing is used between unlike serial number recordings. By such an arrangement, it is less difficult to scan a page and observe excessive repairs for a given instrument.

Such excessive repairs are noted at the time of review for corrective action. Corrective action can take a number of courses:

Complete overhaul by Bendix Radio or equipment manufacturer.

Engineering investigation by Bendix Radio or equipment manufacturer for corrective action.

Disposal of extended service instrumentation.

Instrument History Report

Instrument type	Instrument manufacturer	Instrument model number	Asset - T.O. - G.F.E. - Pool - serial number	Date received	Date delivered	Department	Calibrator's class., plant, shift	Calibrator's clock No.	Total hours of calibration	Disposition	Repairman's class., plant, shift	Repairman's clock No.	Total hours of repair	Repair code #1	Repair code #2	Repair code #3	Repair code #4	Calibration reschedule code
01	AIB	00065376	001494765	113	130	9532	A	1218	02.2	7	B	1272	12.0	CA0024W	RE0038T	TR0004W	VA0003V	A
01	AIB	00065376	001494765	301	310	9531	A	1034	1.8	6								D
05	BEE	00000426	001494767	120		4550	J	4528	1.0	5								C
05	BEE	00000426	001494767	120	121	4550	J	4528	1.2	9	J	1214	2.0					C
05	AIC	00000042	402769101	205	222	4550	B	1209	1.5	9	A	1212	3.5	VA0001X				F
10	AIC	00000046	402769092	310	321	4622	B	1209	1.5	7				RP0042Y	ML0000P			R

The final report to be discussed will be Justification of Purchase of Instrumentation for Standards Laboratory.

Such purchases of instrumentation, in most cases, have been a problem for standards laboratories. This is due primarily to the fact that most standards laboratories are a part of the indirect labor base of a plant accounting system. As a result there is always pressure by management to keep this cost from rising. This, for my money, is just what the standards laboratory manager needs to sell management: purchase of equipment to reduce operational cost.

In the case of excessive repairs the standards laboratory manager can evaluate the repair cost in comparison to average or standard cost of a given instrument over the amortized period and make an intelligent decision on any instrumentation within the Division, including his own equipment.

In regard to standards laboratory purchases, savings can be evaluated in 90 percent of the cases where purchases of instruments are involved. Such justifications can be and are made on the basis of DP information of repair and calibration histories.

Purchases of new instrumentation with increased accuracy is the most significant problem in regard to instrument justification by the standards laboratory management.

Take the General Radio, Type 1615A Capacity Bridge, as an example. The primary purpose of purchase of this instrument is the increase of accuracy of capacity measurements from ±.1 percent to 0.01 percent and capacity range from 10^{-10} fds to 10^{-17} fds.

However, to justify such purchase one needs only to look into this matter to determine, for most purposes, that this instrument is a direct reading device with little need for use of correction charts in applications. Since the instrument is a three terminal device, we can eliminate the guard circuit, which is quite difficult and time consuming to balance.

One must then evaluate this time in the form of labor cost when this Type 1615A bridge is used. The second consideration is to review the type of equipment to be calibrated and evaluate savings. In the case of the General Radio Bridge, the savings were valuated thus:

Calibration of

Quantity	Mfgr.	Type	Nomenclature	Total time old system	Reduction new system	Savings 5 years
28	G.R.	219	Capacitor Decade	33.4 hr./yr	5 hr./yr	$71.50
39	G.R.	722	Prec. Cap.	118.4 hr./yr	59.2 hr./yr	917.62
18	G.R.	716	Capacity Bridge	83.8 hr./yr	21.0 hr./yr	325.50
				Total Savings		$1314.62

At this point it can readily be seen that 87 percent of the capital investment can be recovered on the basis of time study and the data processing history of calibration. This then becomes a factual document. Without data processing history it would become quite a problem to obtain such a history.

In the case of this extended bridge accuracy of the General Radio Type 1615A Bridge, it is possible to certify one of the laboratories' standards for an annual direct labor cost of $68.00, where previously this same capacitor cost $475.00 for NBS certification. As a result of this savings an additional $2035.00 in operating cost can be realized. Total saving represents $3349.62 or a return of 222 percent of the investment.

This may seem fantastic but is nevertheless true. What is important, it can be backed up with factual data processing information.

You may say to yourselves that this is no magic method of justifying capital expenditures. My answer to you is this; there is no easy way to estimate savings. I can tell you this, though: we in the standards laboratory at Bendix Radio have not had a rejection of capital equipment request since the use of data processing. Our management is extremely cost-conscious and difficult to sell if the facts are not properly presented.

10. In Conclusion

History is established in most cases much more cheaply by the use of data processing. Such information stored in any other manner is too costly to obtain. In cases where cost is no object, timing is of great importance, this poor timing has resulted in educated guessing.

The most time-consuming problem of data processing is the initial development of the document: Beyond this point DP becomes an asset for any management. Poor guessing can result in high operating cost and loss of future business due to this excessive operating cost.

You as managers of standards laboratories help to reduce this operating cost by supplying factual information to make decisive decisions and obtain valuable results in terms of efficiency and accuracy.

The future holds great possibilities in storage of information by the use of magnetic tapes and computers. Such information now generally is card programmed. Tape storage offers even greater flexibility and time savings to extract data. Such data is at the disposal of managers to make decisions quickly based on these stored facts of history.

We at Bendix Radio are, at the present time, programming all of the information presented into the IBM 1401 computer. By such programming no mathematical calculations or tabulation will be performed as a hand operation. By so doing, we will reduce the time required to establish a useable report.

Instrument Fault and Calibration Recycle Code

Component	
Attenuator	AT
Battery	BA
Bearing	BE
Ballast Tube	BT
Bolometer	BO
Capacitor	CA
Cavity	CV
Chopper	CH
Crystal (Quartz)	CQ
Diode (Semiconductor)	DI
Dial	DL
Fitting (Coax)	FI
Fuse	FU
Fuse Holder	FH
Generator	GE
Jack	JA
Knob	KN
Lamp	PP
Lamp Socket	PS
Inductor (Choke or Coil)	LC
Meter	ME
Motor	MO
Neon Lamp	NE
Plug	PL
Probe	PB
Potentiometer	PO
Resistor (5 to 20%)	RE
Resistor (1% & Better)	RP
Rectifier (Selenium, Silicon & Oxide)	RS
Relay	RY
Standard Cell	SC
Stepper Switch	SS
Stylus	ST
Switch	SW
Syncro (Resolver)	SY
Tape (Punch)	TP
Terminal	TE
Thermocouple	TH
Transformer	TR
Transistor	QA
Tube socket	TT
Vacuum tube	VA
Wiring	WI
Calibration	CB
Case	CE
Construction & modification	CM
Line calls (no repair made equipment sent to Dept. 953)	LT
Mechanical maintenance	ML
Module	MD
Product under test	PR
Strip (RF, IF, delay line, etc.)	SP
Subunit or Subchassis	SU
Test equip. (improper use of)	TF

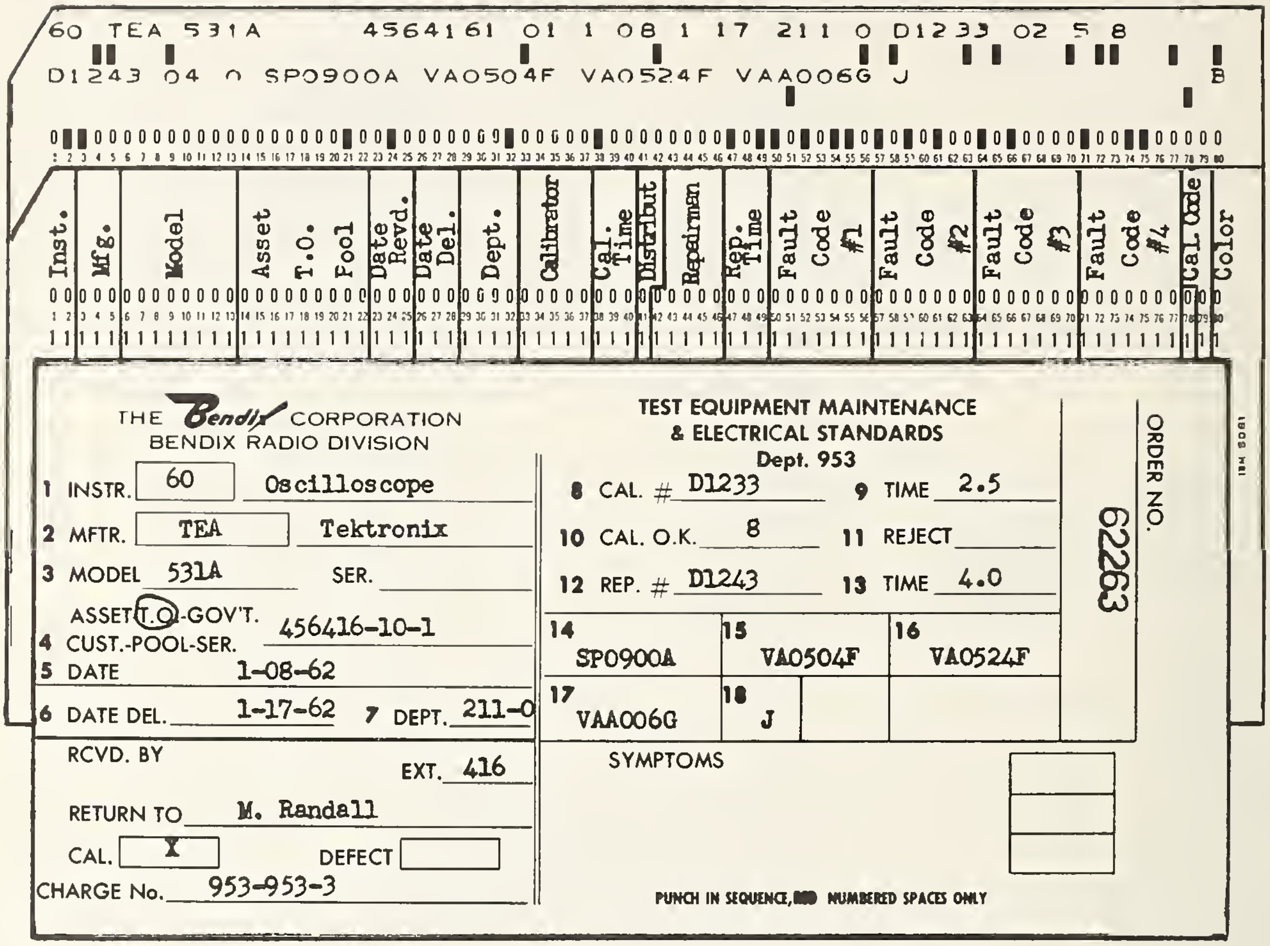
Component Symbol
0001
Use 1st component symbol of a strip when making strip adjustment.

Fault	
Adjustment (electrical)	A
Broken or torn	B
Component (incorrect)	C
Distortion	D
Erratic	E
Frequency response	F
Gain low	G
Gassy	H
High value	J
Hum	K
Intermittent	L
Low value	M
Low d-c output (rectifier)	N
Maintenance--general (Lubrication, new chart paper, re-ink, new pens, new filter, new brushes, etc.)	O
Misaligned (Mechanical adjustment)	P
Missing component	Q
Microphonic	R
No fault found	I
Noisy	S
Open	T
Parts selection	U
Sensitivity or emission	V
Short	W
Stops oscillating	X
Leaky	Y
Worn	Z

A021	Overhaul
B022	General inspection
C006	Excessive repair

Calibration recycle code:

Year	Month	Code	Year	Month	Code
1962	July	A	1963	Aug.	N
	Aug.	B		Sept.	O
	Sept.	C		Oct.	P
	Oct.	D		Nov.	Q
	Nov.	E		Dec.	R
	Dec.	F	1964	Jan.	S
1963	Jan.	G		Feb.	T
	Feb.	H		Mar.	U
	Mar.	I		Apr.	V
	Apr.	J		May	W
	May	K		June	X
	June	L		July	Y
	July	M		Aug.	Z



60 TEA 531A 4564161 01 1 08 1 17 211 0 D1233 02 5 8

D1243 04 0 SP0900A VA0504F VA0524F VAA006G J B

Inst. | Mfg. | Model | Asset T.O. Pool | Date Revd. | Date Del. | Dept. | Calibrator | Cal. Time | Distribut | Repairman | Rep. Time | Fault Code #1 | Fault Code #2 | Fault Code #3 | Fault Code #4 | Cal. Code | Color

THE Bendix CORPORATION
BENDIX RADIO DIVISION

TEST EQUIPMENT MAINTENANCE
& ELECTRICAL STANDARDS
Dept. 953

1 INSTR. 60 Oscilloscope

2 MFTR. TEA Tektronix

3 MODEL 531A SER.

ASSET [✓ T.O.] -GOV'T.

4 CUST.-POOL-SER. 456416-10-1

5 DATE 1-08-62

6 DATE DEL. 1-17-62 7 DEPT. 211-0

8 CAL. # D1233 9 TIME 2.5

10 CAL. O.K. 8 11 REJECT

12 REP. # D1243 13 TIME 4.0

14 SP0900A	15 VA0504F	16 VA0524F
17 VAA006G	18 J	

ORDER NO. 62263

IBM 5081

RCVD. BY

EXT. 416

RETURN TO M. Randall

CAL. X DEFECT

CHARGE No. 953-953-3

SYMPTOMS

PUNCH IN SEQUENCE, [illegible] NUMBERED SPACES ONLY

Exhibit "A"

Exhibit "B" - Calibration Recycle Report

6 Months - Instrument Type - June - December

Instrument type	Instrument	Total cal.	Total rep. & cal.	Total repair	Reject #2 1 Mo.	Reject #4 from repair	Reject #5 over 1 Mo.	$\frac{2}{2+6+7}$ % Reject 1 Mo.	$\frac{2+5}{2+5+6+7}$ % Reject over 1 Mo.	Accept. #6 minor adjustment	Accept. #7 cal. only	Accept. #8 minor adjustment from repair	Accept. #9 from repair	Total units	$\frac{4}{4+8+9}$ % Reject from repair
1	Amplifiers & receivers	138	163	24	15	8	19	12.0	23.7	47	62	23	127	325	5.0
5	Attenuators	306	68	8	18	11	20	6.2*	12.4	11	257	5	52	382	16.1
7	Load, R.F.	9	1	1							9		1	11	
10	Audio oscillators	289	147	11	35	16	36	13.6	24.3	24	197	10	118	447	11.1
15	Bridges	150	78	22	19	8	22	14.7	27.1	76	34	23	46	250	10.3
20	Capacitors, resistors, Inductors	116	14	7	7		3	5.8*	8.1*	27	85	4	4	137	
25	Distortion meters	38	44	4	4	7	4	11.7	21.0	11	19	4	33	86	15.9
30	Frequency meters	322	217	102	61	6	10	19.8	21.6	52	195	13	196	641	2.7
32	Phase indicators	6		1						4	2			7	
38	Transistor & tube Analyzers	13	6	3	3			20.0	20.0	9	3	2	2	22	
39	R.F. Wattmeters & cal. Detectors	127	37	3	10	3	4	7.9*	10.7	32	84	10	21	167	8.8
40	Audio power meters	72	13		1		3	1.4*	5.6*	2	65		14	85	
41	Multimeters	262	397	11	33	51	71	15.4	36.4	53	128	10	313	670	13.6
42	Voltohm meters, V.T.V.M.	628	525	22	61	84	150	12.2	32.6	228	208	91	330	1175	16.6
43	Meters 1% >	120	55	2	9	4	10	6.9*	13.6	5	115	3	29	177	11.1
44	Q Meters	60	37	2	14	6	5	25.4	31.6	1	40	1	30	99	16.2
45	Meters <1%	682	110	3	23	9	35	3.4*	8.2*	24	625	3	72	795	10.7
46	Multipliers & shunts	14	1	2	1		1	7.6*	14.2	1	11		1	17	
48	Digital voltmeters	23	24	32	3	3	5	25.0	47.0	7	2	14	7	79	12.5
51	Test jigs	158	31	58	14	3	10	9.3*	15.0	46	89	3	24	247	10.0
52	Miscellaneous test equip	786	64	1021	14	3	7	1.7*	2.6*	537	235	11	43	1872	5.2
60	Oscilloscopes	600	683	409	104	29	118	24.5	41.0	182	138	212	439	1692	4.2
64	Power supply absolute Voltage	11	7	4	1		1	10.0	18.1	3	6	3	4	22	
65	Power supplies	272	162	93	23	17	27	9.4*	18.4	113	108	8	138	527	10.4
66	Hi pots	27	12	1	4	1	1	15.3	18.5	3	19	1	10	40	8.3
75	Signal generators	701	615	133	113	67	102	19.0	30.9	172	308	156	378	1449	11.1
76	Acceptance bendix	30			5		1	17.2	20.0	8	16			30	
77	Acceptance vendor	524	5	1	53		5	10.2	11.1	39	423	2	7	530	
	Totals	6488	3518	2029	648	336	670	11.0	20.2	1718	3487	612	2440	12036	9.9

▭ Above acceptable target.
* Below acceptable target.

Exhibit "C" - Work Load Forecast Report

Instrument Type Code	Instrument	Cost per unit	Cost per unit per year	Calibration period in months	Cycle per year	Average time/unit repair and cal.	Active equipment units	Hr/yr ÷ by 1822.5 hr 1 man year in hr of work
1	Amplifiers & receivers	22.40	44.81		2	3.91	355	1.523
2	Recorders	-	-		-	0	61	-
5	Attenuators	9.33	9.33	12	1	1.629	380	.339
7	Load, R.F.	7.84	7.84	12	1	1.368	51	.038
10	Audio oscillators	14.90	89.42	2	6	2.601	367	3.143
15	Bridges	21.82	21.82	12	1	3.808	115	.428
16	Comparators	21.82	87.28	3	4	3.808	203	1.697
20	Capacitors, resistors, Ind.	9.20	9.20	12	1	1.606	559	.493
25	Distortion meters	17.41	69.63	3	4	3.038	115	.767
30	Frequency meters	17.22	103.35	2	6	3.005	385	3.810
31	Absorption type frequency meters	17.22	34.44	6	2	3.005	104	.343
32	Phase indicators	13.88	55.54	3	4	2.423	42	.223
38	Transistor & tube analyzers	17.57	35.14	6	2	3.066	58	.195
39	R.F. wattmeters & cal detectors	14.00	56.02	3	4	2.444	91	.488
40	Audio power meters	22.00	29.26	9	1.33	3.84	242	.678
41	Multimeters	7.78	15.46	6	2	1.349	1219	1.805
42	Voltohm meters, V.T.V.M.	12.91	51.64	3	4	2.253	82	.405
43	Meters 1% >	3.61	4.80	9	1.33	.630	326	.150
44	Q meters	16.07	96.44	2	6	2.805	62	.573
45	Meters < 1%	3.04	6.07	6	2	.530	1468	.855
46	Multipliers & shunts	5.63	11.27	6	2	.983	43	.046
47	Radiation counters	5.73	11.46	6	2	1.0	26	.029
48	Digital voltmeters	31.64	131.10	3	4	5.522	17	.213
51	Test Jigs	10.46	25.11	-	2.4	1.826	206	.495
52	Cal. special bendix test equip.	8.35	20.06	-	2.4	1.458	825	1.584
53	Uncal. special bendix test equip.	6.42	-	-	2.4	1.121	778	1.148
60	Oscilloscopes	18.12	108.71	2	6	3.162	553	5.757
62	Spectrum analyzers	18.12	36.24	6	2	3.162	309	1.072
64	Power supply absolute voltage	25.50	51.00	6	2	4.450	27	.132
65	Power supplies	15.07	30.14	6	2	2.630	893	2.577
66	Hi pots	7.78	10.35	9	1.33	1.358	25	.025
70	Signal generators	23.48	140.89	2	6	4.098	699	9.430
72	Sweep generators	23.48	46.96		2	4.098	99	.445
	Item #	8	9	1	2	3	4	7

Total manpower requirement = 40.48.

Session 7. Calibration Recycle Analysis and Work Load Control

PANEL ON CALIBRATION RECYCLE ANALYSIS AND WORK LOAD CONTROL

Moderator:
Jerry L. Hayes, Bureau of Naval Weapons
Pomona, California

Members:
Joseph M. Aldrich, Ryan Aeronautical Company
San Diego, California

John W. Rodgers, Bendix Radio Division
Baltimore, Maryland

Michael Rothbart, National Astro Laboratories
Pasadena, California

John R. Van de Houten, Aerojet-General Corporation
Sacramento, California

Andrew J. Woodington, guest, General Dynamics/Astronautics
San Diego, California

Summary

Question: "Has any work been done on establishing a useful shelf life for a calibrated instrument with a twelve-weeks' calibration interval, which may lie on the shelf for ten weeks before check-out from the instrument pool? How do you determine the date that recalibration is due?"

Mr. Woodington stated that his organization manages to keep two items in current calibration. If there are five items on the shelf, only two are kept in calibration, the rest being kept in temporary storage. There is no purpose in having all of the items in the crib cycling through the laboratory for recalibration.

Mr. Van de Houten said that they use a great number of transducers in their testing operations, and for a period of years they have plotted shifts that occur in usage and shifts that occur during shelf life. They have found no significant shift in calibration due to storage.

In answer to a question by the Moderator, Mr. Van de Houten replied that their data shows that the instrument doesn't age while on the shelf. However, his answer referred only to certain types of transducers and not to electronic equipment.

In contrast, the Moderator said that in his experience altimeters left on the shelf were in worse shape than if they had been used because they developed flat spots.

Question: "Can the instrument manufacturer suggest or estimate suitable calibration intervals for his instruments for the average engineering and production application?"

The Moderator said that if the instrument manufacturer would supply a good calibration procedure with his instrument, he could then estimate a calibration interval. This is done in military procedures now.

Question: "Should NBS have a policy of mandatory recall? If 'Yes', why? If 'No', why? Should other standards laboratories assume responsibility and control?"

In reply, Mr. Rodgers throught Dr. Astin had indicated earlier that there was some question as to the necessity for a mandatory recall. The Bureau has published *Suggested Practices for Electrical Standardizing Laboratories,* [NBS Circular 578] which outlines basically what these requirements should be. However he believed it is entirely up to the practices of the local corporation to decide at what intervals a calibration is necessary, based on their particular use. If they have one reference standard that is used once every six months and is subject to proper handling in normal use, it should not be necessary to return this standard to NBS every six months or every year. He pointed out that the publication mentioned is directed more to reference standards than to the calibration of general-use equipment.

Mr. Woodington said that, fundamentally, the premise is that NBS is not involved contractually in your business. Therefore they have no reason to have a recall system, whereas it is a corporation's own problem to see that the procedures are correct. The reason you have a recall system in your own corporation is because of the contracts you have. You are supporting an end-item product. He commented further, "At this time, I will try to turn the triangle around. Most people list the calibration triangle with the standards laboratory on top and the production end on the bottom. I say turn it over, because if you ever lose the production items you won't have a standards laboratory."

Question: "When in-place calibration practice is used, is there any good plan which is recognized for preventive maintenance control; or can maintenance and calibration be separated?"

Mr. Rodgers said that he didn't quite understand the difference between preventive maintenance and a routine cycle of calibration, that in his opinion they are one and the same, and that preventive maintenance would obviously entail the corrective action as a result of calibration.

Mr. Woodington concurred with Mr. Rodgers, and said that he didn't understand the difference either.

Mr. Van de Houten said that you have redundant calibration if you have both in-place calibration and preventive maintenance.

Question: "I believe some standards laboratories overlook a very important operation in the calibration check on a given instrument. This is the as-found or as-received check. Would the panel please comment on the importance of such a check on the incoming instrument, prior to disturbing the as-received condition?"

Mr. Rothbart said that he thought it important to find out whether an instrument is in calibration when received and that it should be noted on the record when specific data are taken.

The Moderator commented that if the policy is to check the instrument to specifications, which is what 90 percent of the working level calibration laboratories do when they test equipment, then it would be significant to find out what the original condition of the instrument is before adjusting it. He felt that if it is within specifications, it should not be readjusted.

Mr. Rodgers said that the equipment manufacturers have a specific tolerance for their instruments, and the user of the instrument should expect that it will remain within these limits after calibration. His company has data to show that if the instrumentation is readjusted to nominal, it may go out of tolerance after it leaves the laboratory.

Mr. Woodington added that his people are instructed that in most cases the manufacturer will set an instrument before it leaves the factory to within one-half of the tolerance that he guarantees. This allows some drift or movement during the ninety-day warranty. Astronautics has decided to use this same policy. Therefore their people are instructed that if the instrument is within one-half of its advertised manufactured tolerance, they are to do nothing to it, and they must state on the records why they decide to make any adjustment. This is taking a somewhat pessimistic view in that, if during a calibration interval the item doesn't drift more than halfway to the limit of the manufacturer's tolerances, it will remain within tolerance during the next calibration interval. But there seemed no reason for giving the user an instrument that is on the borderline, because then there is no room for drift during a calibration interval. One can find examples where, for some reason or other, an instrument may drift backward toward the nominal value, but you can't depend on this.

Mr. Rodgers commented that they work to 90 percent of tolerance rather than to 50 percent and have found this quite successful.

Question: "Do most standards laboratories report calibration data quantitatively or qualitatively?"

Mr. Rodgers answered by saying his company does not record data unless data processing feedback indicates they are in specific trouble with specific instruments. Their philosophy is that the recording of data is no better than the people who record the data. Once you record the data, what do you do with it all? If it is placed in a file cabinet and nothing done with it, little is gained. When their data processing system indicates that any particular instrument is giving exceptional trouble, they will evaluate it with the use of specific data. During interim periods or in the normal calibration of instruments within the standards laboratory, they determine the quality of work that the employees are producing by quality control methods of sampling the daily outputs.

Mr. Van de Houten cited some of the answers that came back on the NCSL Committee survey. He reported that the number of people recording data is higher than ordinarily would be expected. Out of the 72 replies, 34 (or 47.2 percent) answered that actual data are being recorded using preprinted forms showing each test point. Another 20 (or 27.8 percent) reported that the actual data are recorded without use of special forms, and reliance is placed upon the technician to determine test points and tolerances. Only 10 out of the 72 who said that recording is not done and is not planned for the future. He was certain there was some bias; for example, some calibration people are recording full data while other laboratories in the same company are recording only partial data.

Mr. Rothbart said that in providing calibration services for the Aerospace Industry, they find that the prime contractors in the military take data and require it. They find in going through the various tiers of subcontractors that, for the smaller companies, go and no-go data seems adequate. But they have noticed in the past two years that the trend very definitely has been away from the go and no-go data to the recording of actual measurements.

Mr. Woodington maintained that data are absolutely necessary. However, he does not put out correction curves, which are often requested. He raised the question, "How can you predict what course your standards are taking if you don't have some data?" This information is not apparent immediately. "We know that standards are not invariant, and although I don't recommend adjusting your secondary standards all the time, you should have the data to predict where they are going to go. We have found, in examining the data we get back from NBS, that we can predict about when a particular item is going to go out of the manufacturer's tolerance. This tells us, incidentally, that we have to make up our minds whether we are going to continue its use or to buy a new one. This may be a shock to some of the manufacturers of instruments, but we do have to replace their items once in a while."

At the Moderator's request to comment on the difference in cost between not taking data and taking data, Mr. Woodington replied that he had been forced into data keeping [referring to the adjustment of instruments and not to the calibration of transfer or reference standards]. He found that it was worthwhile and that it required but 3-1/2 minutes, on the average, of a technician's time to record the data.

Responses to additional questions directed specifically to Mr. Rodgers were submitted later, as follows:

Question: "Your workload-manpower forecast was based on 40 productive hours/week. What is the actual efficiency of the technicians in your laboratory, i.e., what are the available hours vs. hours actually spent in calibration and repair?"

"Our manhours are based on 7.5 hours per day or 37.5 hours per week, rest period, cleanup time, etc., not included. Since most repair and calibration work is covered by written procedures and is more or less standardized, the personnel engaged in routine repair and calibration must account for 7.5 hours of work per day. Absence, training, etc., represent approximately 5 percent of the total manpower and are included in all forecasting."

Question: "How do you fix a standard time on repair hours when there is almost an infinite number of failures or combinations of failure modes?"

"Repair time standards represent averages and are based in most cases on a large volume of work load and/or cycles in a given category. These averages sometimes are questionable when small sampling is used. Such methods are valid only when large sampling is evaluated. Since our volume is in excess of 2000 instruments per month, the small questionable samples do little to reflect errors in the work load and manpower forecasting."

Question: "Are details of ADP [Automatic Data Processing] available from Bendix?"

"The written paper [presented by Mr. Rodgers] gives a detailed example of reports for the ADP operation. If the question is directed to the programming of the data processing machines, I am sure that any qualified machine programmer can convert the information from the work card to a final report. Detailed programming of ADP equipment is guided by the equipment available at a specific location and therefore cannot be standardized."

Question: "What is the labor cost per calibration of instruments, or what is the labor hours ratio to calibrations per month at Bendix?"

"There is some uncertainty as to what is meant by this question. I am assuming that it means average time for the total calibration work processed. This time is 1.2 hours per calibration. It includes all work processed by the laboratory of working test equipment. Calibration of reference standards, special investigations, acceptance testing, etc., are excluded."

Session 8. NCSL Relations to and Cooperation with Technical Societies

PANEL ON NCSL RELATIONS TO AND COOPERATION WITH TECHNICAL SOCIETIES

Moderator:
William Amey, Leeds and Northrup Company
Philadelphia, Pennsylvania

Members:
Arnold W. Young, representing AOA
Orval L. Linebrink, representing ISA
Thomas A. Marshall, Jr., representing ASTM
Ray E. Davis, representing ASQC
Harvey W. Lance, representing IRE

Summary

Mr. Amey: In the process of growing and working out ways in which it can be of service to the nation, the NCSL recognizes the possibility that its efforts may duplicate the efforts of other groups. In many places today there is concern over the optimum usage of technical personnel, and it is a deplorable situation in which a few people find themselves at one moment representing one society on a certain subject and the next moment dashing off to another meeting a thousand miles away to represent another organization on essentially the same subject. When there is no common meeting ground, NCSL must recognize the fact and act accordingly, but when there is a common meeting ground NCSL should make every effort to cooperate with other societies in order to make optimum usage of available resources. We now have reached the point at which we can define some of our problems and refer them to other societies to determine their interest in cooperating with us. Not all of the societies with which we must make contact are represented on the panel.

Mr. Young: AOA (The American Ordnance Association) has only one word to justify its existence, and that word is "preparedness." The Association is a national society concerned mainly with military activities in the United States, and many of its members are from standards laboratories. Of specific interest to NCSL is the Standards and Metrology Division, which has dimensional, electronic, and physical groups. Annual meetings held by this Division consist normally of workshop discussions, more formal presentations to the whole meeting, and usually tours of a facility of interest to the attendees. There are a lot of familiar people at the Standards Laboratory Conference. This indicates both a continuing desire on the part of the conferees to keep well informed and a common field of interest between the National Conference of Standards Laboratories and other groups. The basic program of NCSL is believed to be the dissemination of information. This is a real need. To fulfill this need will require a full time staff or else affiliation with some other organization which is well staffed.

Mr. Linebrink: ISA (The Instrument Society of America) has undergone a rapid development over the past 15 or 16 years. Of most direct interest to NCSL is the Measurement Standards Instrumentation Division. That division was formed within the last three years. At the ISA meeting a year ago in Los Angeles there were six scheduled sessions on standards with an average attendance of about 150 per session, and 10 sessions are scheduled for the New York meeting later this fall. The interests of this Division are in the laboratory applications of measurement, contrasted with standards of practice which are covered by another ISA division. The interests of the Measurement Standards Instrumentation Division are somewhat broader than just the standards laboratory. ISA through this Division has many common interests with NCSL and in fact was the host to the Ad Hoc Committee on several occasions preceding this conference. There was coordination in the scheduling of the Standards Laboratory Conference here in August and the

ISA meeting in October in New York. This and other types of cooperation are desirable to avoid unnecessary overlap and duplication. There is a need for better communications among the technical societies. In many cases, instrumentation and measurement standards can be a common denominator for improved communication. ISA offers to consider any feasible steps for cooperation with NCSL.

Mr. Marshall: ASTM (The American Society for Testing and Materials) was formed in 1898. Its purpose is the promotion of knowledge of the materials of engineering and the standardization of specifications and methods of testing. ASTM has four major classes of membership: personal members, institutional members, and industrial members who may be either sustaining members or just straight company members. The technical activities of ASTM are conducted by 89 technical committees having a large number of members, including people who are not members of the society, joint committees with other societies, and special committees working closely with ASA. Committee E-1, on methods of testing, is perhaps of special interest to NCSL. ASTM offers its cooperation in any way that would promote the technology of our country, including working jointly with NCSL and providing services whenever possible.

Mr. Davis: ASQC (The American Society for Quality Control) consists of members whose interests are in a quality product. The society has lots of divisions and many technical committees. The National Metrological Technical Committee has interests closely paralleling the interests of NCSL. Several of the topics listed in the Suggested Practices Manual are examples of topics in which ASQC also is interested. ASQC is having a conference in Denver in October which will cover some of these topics of mutual interest. ASQC is interested in the many big companies which are well established, as well as the many new ones which are established each year and which are most likely to need help. The majority of the interests of NCSL seem to be aimed at the larger and well established companies. NCSL should consider seriously whether it wants to remain independent or to affiliate with some other organization. ASQC offers whatever assistance it may be able to give. Without enveloping NCSL, it would like to join NCSL in a common endeavor.

Mr. Lance: IRE (The Institute of Radio Engineers) is an individual membership society, in contrast to NCSL, which is primarily laboratory oriented. There are now 29 IRE Professional Groups, which are divisional-type organizations having, in general, work directed toward a particular subject matter area. The Professional Group on Instrumentation (which is in the process of changing its name to the Professional Group on Instrumentation and Measurement) is a co-sponsor of next week's International Conference on Precision Electromagnetic Measurements. There is also a committee structure within IRE which does quite a lot of the work of the Institute. For example, there is a Standards Committee, the main work of which is to review and approve standards originating throughout the organization. There also is a Measurement and Instrumentation Committee, which has a Subcommittee on Basic Standards and Calibration Methods. This subcommittee is engaged in a rather ambitious project, the goal of which is to get a consistent statement of what is available and what is needed in the way of standards and measurement accuracy at various levels. There is relatively little attempt on the part of the organization to guide and direct the efforts of its members, and the output of the Institute reflects the integrated interests and work of the individual members. However, information on many of the needs noted at this conference could very well be fed to the IRE members through the professional group structure or through the committee structures and thus could be effective in guiding work of the Institute. The IRE would welcome the opportunity to be of assistance in the field of electronics in meeting technical needs identified by the NCSL.

The AIEE (American Institute of Electrical Engineers) was represented not in person, but by a letter from Dr. B. R. Teare, Jr., President of AIEE. Dr. Teare noted the forthcoming merger of the AIEE and the IRE and expressed the opinion that the merger would simplify the problem of liaison between the merged society and NCSL. He referred to the close connection already existing between NCSL and AIEE, by noting that many of the people active in NCSL also have been active on AIEE committees engaged in the formulation of electrical standards and also have participated actively in technical conferences sponsored by AIEE. He recognized that the interests of NCSL are broader than electronics and electricity, but in that field he assured NCSL of any feasible measure of cooperation that might be suggested. This offer included assistance by individual AIEE members, assistance by the technical committees and, in particular, assistance by the standards committee itself. He offered the use of various channels of communication at the disposal of AIEE in publicizing programs initiated by NCSL on effective use of precision methods and equipment. He stated that the monthly magazine Electrical Engineering is available for assistance, noted that help could be had through the committee organization, and commented finally that the Instrumentation Division would be delighted to carry and support the message of NCSL.

Mr. Amey: It is most striking that there is a common denominator of recognition, by the societies represented, of the specific needs of the laboratories represented by NCSL, and simultaneously there is a strong offer to help in any way that NCSL thinks advisable and that is feasible for those societies to provide. It appears that in order to get the cooperation of the societies, all we have to do is to define our problem so that it can be understood. The organizations represented here today are by no means all of the organizations with which NCSL would like to maintain liaison. For example, there are the Precision Measurements Association and the Precision Measurements Society. At this point nobody in NCSL is advocating that we join up with anybody else. However, there is a strong feeling that NCSL should encourage other organizations to work on problems it identifies and that NCSL itself should undertake only tasks that cannot or will not be done satisfactorily by other organizations.

NBS Miscellaneous Publication 248
Proceedings of the 1962 Standards Laboratory Conference

Session 9. Recommended Practices for Standards Laboratories

PANEL ON RECOMMENDED PRACTICES FOR STANDARDS LABORATORIES

Moderator:
K. G. Overbury, Sandia Corporation
Albuquerque, New Mexico

Members:
M. J. Leight, Hughes Aircraft Company
Culver City, California

Mary E. Hoskins, Sheffield Corporation
Dayton, Ohio

Frank D. Weaver, National Bureau of Standards
Boulder, Colorado

Robert G. Davison, USAF Calibration Division
Newark, Ohio

Jerry L. Hayes, Bureau of Naval Weapons
Pomona, California

Summary

Mr. Overbury: A preliminary version of a manual entitled, "Suggested Practices for Standards Laboratories," has been prepared by the Committee on Recommended Practices. Several of the special committees of NCSL did the initial work in preparation of material for the various subject matter areas. No detailed guidelines were provided. Hence the material is of a general nature and mostly is still in outline form. The manual is intended to be generally applicable to various types of standards laboratories, and the using laboratory may choose those practices which it wishes to follow. The committee considered including in the manual a glossary of technical terms, but time did not permit making such a compilation for the preliminary draft.

Mr. Leight: A standard practices manual would be valuable in achieving uniformity among the divisional standards laboratories of a corporation. It would be helpful in establishing new facilities, in planning future work for the laboratories, and in budgeting for standards laboratory operation. It also would be helpful in comparing one's own laboratory with the laboratories of other companies and in evaluating the capabilities of suppliers and subcontractors. For widest applicability the practices would need to be stated rather generally and broadly so that they could be flexibly applied.

Mrs. Hoskins: There is need for better communications within and among standards laboratories, and many benefits could be achieved from the use of a common terminology. The Suggested Practices Manual should be quite helpful along this line. It should be useful in connection with the internal evaluation of capabilities and weaknesses, it should help to resolve conflicts between buyer and seller, and it should help others to tell whether the services of a particular laboratory would likely be helpful to them.

Mr. Weaver: A recommended practices manual would be helpful to both new laboratories and to old ones. The common language and common outlook provided by such a manual would enable other laboratories to do work of greater uniformity and higher quality and thus would make a valuable contribution toward the general improvement of measurement accuracy. An NCSL newsletter also has great potential value because it can provide information which is more up-to-date than that in the Suggested Practices Manual.

Mr. Davison: The Suggested Practices Manual can form the basis of self-imposed discipline for standards laboratories. This is good, because discipline is less objectionable if it is self-imposed. The manual should result in a general improvement in the ability of contractors to meet government contractual obligations.

Mr. Hayes: The preceding comments on the matter of self-discipline are very much to the point. A good Suggested Practices Manual, properly followed, would result in more freedom for the operating laboratories while at the same time improving the quality of the company's product. The Navy is interested in the National Conference of Standards Laboratories because of the expected improvement in self-discipline. Sometime ago the Navy had to issue its own suggested practices manual. If NCSL had been in existence earlier, this effort on the part of the Navy might have been unnecessary. The military standards laboratories have the same types of problems as those encountered in industrial standards laboratories, and the Suggested Practices Manual should be of value in the operation of these laboratories.

Discussion Period

Question. "What is the difference between a 'standards laboratory' and 'calibration laboratory'?"

Often one can't tell them apart. When a distinction is made, the term "calibration laboratory" usually is applied to the laboratory that calibrates test equipment and measuring instruments used in the shop or on the production line, while the term "standards laboratory" is applied to the laboratory that calibrates standards for the calibration laboratory.

Question. "Would the military endorse the Suggested Practices Manual by making it a part of a military specification?"

It is unlikely that the manual will become a part of formal specifications, but it may become recognized and referenced as a useful means of achieving the measurement ability required to fulfill military contracts. It should be recognized that the suggestions of a recognized authority carry considerable weight. Hence, looking toward the future, NCSL is shouldering a great responsibility in undertaking the preparation of such a manual.

Question. "From an industry point of view, how can we be absolutely certain that the suggested practices are broad enough to cover all types of needs?"

It is doubtful that they can be made sufficiently broad to cover every type of need. Hence, it is unlikely that the manual will become an ironclad list with which industry is required to comply. However, if the suggested practices fit the operation economically and technically, then there is no reason why they should not be used.

Question. "Should equipment and distribution centers commonly referred to as 'cribs', and equipment utilized in quality control engineering laboratories and manufacturing areas, be under the direction of a calibration laboratory?"

That depends on where the calibration laboratory fits into the organization. If it serves a line function, it has one type of responsibility. If it serves a staff function, it has a different type of responsibility. Some companies have the calibration laboratory directly under quality control and others do not. In a survey by the NCSL Workload Control Committee it appeared that the choice depends greatly on the size of the company and on how many calibration laboratories there are in the company. About half of the companies surveyed had loan pools in their calibration or standards laboratories, and the other half had what they called department loan pools.

Question. "Taking into consideration the present pace in the field of electronics, space, and military weapons, could you give some indication as to how these suggested practices can be kept up-to-date and valid? Also, when will the next draft of the manual be issued?"

There is no definite target date for the next draft. Since the manual is the result of a committee effort, perhaps subcommittees covering different geographical areas should be established and problem-oriented workshops should be held in each area in order to speed up the preparation of the manual. This type of committee operation also could be used to keep it up to date.

Question. "In view of the large potential use of Army, Navy, and Air Force calibration procedures and of their high cost in bulk, what are the chances of making them available on microfilm?"

The feasibility of using microfilm depends upon how many copies are wanted. One company tried microfilm, and the only result was headaches and inefficiency. Some of these procedures are available for purchase through the Office of Technical Services of the Department of Commerce. Possibly that office could supply reproducible copies for those having need of a large number of copies of any one procedure.

Question. "Has any clearing house as such been established in the armed services to standardize calibration procedures for common equipment, including such factors as parameters to be measured, degree of accuracy, and calibration interval?"

The Army, Navy, and Air Force have a calibration conference which meets periodically to discuss such problems, and they have a working group on calibration procedure consolidation. This working group has been hard at work on the problem, but it is difficult to solve because the three Services have different equipment and different requirements.

Question. "What decisions were made on standards laboratory information dissemination during the Wednesday evening session?"

That session was more of a progress report. However, the subjects discussed tie in very closely with this session. For example, an index of procedures might well come from a measurement information center instead of being part of a suggested practices handbook. Possibly also the Suggested Practices Manual should be compiled as a series of separate documents, and these might be issued from the information center.

NCSL has started negotiations with NBS to see whether NBS will supply NCSL with a formal secretariat service. If this is provided, it should be possible to speed up work on some of these matters.

Question. "How will it be possible to write recommended practices and procedures prior to clearly defining terminology? In order to achieve standard practices, must we not all talk the same language?"

A common understanding of words and terminology is basic to the entire field of precision measurements. The importance of an adequate glossary cannot be overstated. Yet no progress will ever be made if all we do is get together occasionally and agree that the task is difficult. The experience of other organizations should be called upon in the preparation of such a glossary. The opinions and assistance of a large number of people will be required in compiling, reviewing, polishing, editing, screening, approving, etc. It can easily take two to five years to complete such an undertaking. This statement applies to the glossary and also to the practices themselves.

Question. "To what extent has NCSL considered the formation of area committees, say in Boston, Philadelphia, Washington, etc., to work on problems of interest to the organization?"

This sort of thing was mentioned at the meeting last night. We already have had workshops in Boulder and in Washington, but the suggestion was made of workshops in numerous areas attended only by people within the particular region, say within a 500-mile radius. In previous workshops the attendance was perhaps a little large for effective operation. The regional approach might effectively limit the attendance so that the group could be more effective.

Question. "In connection with measurement audits, would not the presence of an outside supervisor in a standards laboratory prevent the audit from being effective as a measure of the degree of accuracy to which the laboratory normally works?"

Section 5.3 of the manual listed four different types of comparisons: internal, external, personally conducted, and round robins. There are also two kinds of audits. In one the standards laboratory knows it is getting an audit package. The package may be sent in with or without an audit supervisor. The measurement results in this case will be an indication of what the laboratory does under the very best conditions. In another type of audit the standards to be measured are sent to the laboratory in such a way that they cannot be recognized as a part of an audit package. In this case the results are more nearly typical of the day-to-day work of the standards laboratory. One of the best forms of audits can be initiated by the laboratories themselves. The lower echelon laboratory should always make its own cross checks on the calibrations provided by the higher echelon laboratory. Such an effort can be very effective in improving measurement accuracy.

Question. "To what extent will NCSL provide administration in connection with external comparisons?"

It is intended that NCSL provide some small group to administer such comparisons. Its function will be primarily to get the results of the comparison together and distribute them in a confidential manner when necessary. The total results probably should be made available to all participants, but the .results obtained by an individual laboratory would be identified only to the laboratory making the measurements. In any measurement agreement program it is quite important to have a referee. As the program grows to include more and more companies, the necessity of a neutral referee becomes more important.

Question. "In determining whether a contractor is in compliance with military specifications, what sort of documented information do the government inspectors follow?"

The inspectors have manuals and instructions to follow, but there are still problems of communication and of interpreting the meaning of words. The Suggested Practices Manual could be used, when completed, as a guide for government inspectors and by many others interested in measurement requirements and how to meet them.

Question. "How scientifically have existing calibration intervals been determined, and what steps are being taken to find out whether these intervals are proper?"

In one instance, in which 180-day calibration intervals are being used, the interval is an initial estimate for use until better information can be obtained. Probably not enough has been done about readjusting the intervals. More engineering studies are needed. Many factors determine the proper interval. These include the configuration of the instrument, its built-in accuracy and characteristics, and the way it is used. It makes a difference whether the instrument is used one hour a day or eight hours a day, whether it is used on an airplane, on a ship, or in an environmentally-controlled shop. In a small organization it might be realistic to let the technician who has a complete file on a given instrument decide whether the calibration interval should be changed. In a large organization a more formal recall system probably is required.

Question. "Should not the calibration interval be dependent upon the previous history of a standard, and do all laboratories keep a history on the instruments they calibrate?"

According to information obtained by the Workload Control Committee, there are a significant number of laboratories that do not keep a history. There is an approximately equal number that do. In some of the larger companies, the calibration interval is determined by previous histories. This requires quite a paper operation but is believed worthwhile. Some companies have equipment that is standardized every day, some that is standardized every week, and so on up to every five or ten years--all depending on the importance of the standards and the way they are used.

Question. "Should the Recommended Practices Handbook show the paths of traceability to NBS for various levels of laboratories and various types of calibrations?"

The purpose of traceability is to maintain known accuracy at the point of measurement. There usually are many alternative paths or ways of achieving this accuracy. Part of the answer would be for laboratories to establish flow charts so that when questions arise on traceability, they are always in a position to say, "This is the way we are traceable in each of our technologies. This is the way we have achieved the accuracy we claim."

U. S. GOVERNMENT PRINTING OFFICE: 1963 O - 696-881

U. S. DEPARTMENT OF COMMERCE
Luther H. Hodges, *Secretary*

NATIONAL BUREAU OF STANDARDS
A. V. Astin, *Director*

THE NATIONAL BUREAU OF STANDARDS

The scope of activities of the National Bureau of Standards at its major laboratories in Washington, D.C., and Boulder, Colorado, is suggested in the following listing of the divisions and sections engaged in technical work. In general, each section carries out specialized research, development, and engineering in the field indicated by its title. A brief description of the activities, and of the resultant publications, appears on the inside of the front cover.

WASHINGTON, D. C.

Electricity. Resistance and Reactance. Electrochemistry. Electrical Instruments. Magnetic Measurements. Dielectrics. High Voltage. Absolute Electrical Measurements.

Metrology. Photometry and Colorimetry. Refractometry. Photographic Research. Length. Engineering Metrology. Mass and Volume.

Heat. Temperature Physics. Heat Measurements. Cryogenic Physics. Equation of State. Statistical Physics.

Radiation Physics. X-ray. Radioactivity. Radiation Theory. High Energy Radiation. Radiological Equipment. Nucleonic Instrumentation. Neutron Physics.

Analytical and Inorganic Chemistry. Pure Substances. Spectrochemistry. Solution Chemistry. Standard Reference Materials. Applied Analytical Research. Crystal Chemistry.

Mechanics. Sound. Pressure and Vacuum. Fluid Mechanics. Engineering Mechanics. Rheology. Combustion Controls.

Polymers. Macromolecules: Synthesis and Structure. Polymer Chemistry. Polymer Physics. Polymer Characterization. Polymer Evaluation and Testing. Applied Polymer Standards and Research. Dental Research.

Metallurgy. Engineering Metallurgy. Metal Reactions. Metal Physics. Electrolysis and Metal Deposition.

Inorganic Solids. Engineering Ceramics. Glass. Solid State Chemistry. Crystal Growth. Physical Properties. Crystallography.

Building Research. Structural Engineering. Fire Research. Mechanical Systems. Organic Building Materials. Codes and Safety Standards. Heat Transfer. Inorganic Building Materials. Metallic Building Materials.

Applied Mathematics. Numerical Analysis. Computation. Statistical Engineering. Mathematical Physics. Operations Research.

Data Processing Systems. Components and Techniques. Computer Technology. Measurements Automation. Engineering Applications. Systems Analysis.

Atomic Physics. Spectroscopy. Infrared Spectroscopy. Far Ultraviolet Physics. Solid State Physics. Electron Physics. Atomic Physics. Plasma Spectroscopy.

Instrumentation. Engineering Electronics. Electron Devices. Electronic Instrumentation. Mechanical Instruments. Basic Instrumentation.

Physical Chemistry. Thermochemistry. Surface Chemistry. Organic Chemistry. Molecular Spectroscopy. Elementary Processes. Mass Spectrometry. Photochemistry and Radiation Chemistry.

Office of Weights and Measures.

BOULDER, COLO.

CRYOGENIC ENGINEERING LABORATORY

Cryogenic Processes. Cryogenic Properties of Solids. Cryogenic Technical Services. Properties of Cryogenic Fluids.

CENTRAL RADIO PROPAGATION LABORATORY

Ionosphere Research and Propagation. Low Frequency and Very Low Frequency Research. Ionosphere Research. Prediction Services. Sun-Earth Relationships. Field Engineering. Radio Warning Services. Vertical Soundings Research.

Troposphere and Space Telecommunications. Data Reduction Instrumentation. Radio Noise. Tropospheric Measurements. Tropospheric Analysis. Spectrum Utilization Research. Radio-Meteorology. Lower Atmosphere Physics.

Radio Systems. Applied Electromagnetic Theory. High Frequency and Very High Frequency Research. Frequency Utilization. Modulation Research. Antenna Research. Radiodetermination.

Upper Atmosphere and Space Physics. Upper Atmosphere and Plasma Physics. High Latitude Ionosphere Physics. Ionosphere and Exosphere Scatter. Airglow and Aurora. Ionospheric Radio Astronomy.

RADIO STANDARDS LABORATORY

Radio Standards Physics. Frequency and Time Disseminations. Radio and Microwave Materials. Atomic Frequency and Time-Interval Standards. Radio Plasma. Microwave Physics.

Radio Standards Engineering. High Frequency Electrical Standards. High Frequency Calibration Services. High Frequency Impedance Standards. Microwave Calibration Services. Microwave Circuit Standards. Low Frequency Calibration Services.

Joint Institute for Laboratory Astrophysics - NBS Group (Univ. of Colo.).