Traceability: An Evolving Concept

The concept of traceability to national measurement standards has evolved considerably over the past 25 to 30 years. The 1980 paper, *Traceability: An Evolving Concept,* by Brian Belanger [1], is often cited by other authors writing on this subject. This paper and others by NBS/NIST authors have shown why the concept of traceability is not meaningful unless it incorporates quantified measurement uncertainty. Traditionally, measurement traceability has been described as an unbroken chain of comparisons to a primary or reference standard as maintained by a national measurement institute (NMI)—NIST or its counterpart institutions in other countries. The requirement that measurements be traceable to NIST or some other NMI was a response to needs for ensuring consistency among measurements made in different countries or domestic locations (e.g., a liter of gasoline in California should be the same as a liter of gasoline in New York), enabling interchangeability of parts, and facilitating industrial quality control.

The concept of establishing measurement standards and relating measurement results to those of a "higher authority" has been recognized for centuries. When the pyramids were built in Egypt, the length of the pharaoh's forearm was defined as the cubit, and measurements made for its construction were based on that national standard of length. Today, we have fundamental principles of physics, rather than the pharaoh, to rely on for length standards, but the principle is the same—having national standards for key measurements upon which everyone agrees (ideally tracing to internationally agreed upon standards).

Fig. 1. Charles Ehrlich, Brian Belanger, Stanley Rasberry, and Ernest Garner (left to right).

In 1962, the U.S. Department of Defense (DOD) impacted significantly on the evolution of traceability by requiring traceability for DOD laboratories, contractors, and subcontractors in MIL STD 45662A. All measuring and test equipment had to be calibrated utilizing reference standards whose calibration was certified as traceable to the National Bureau of Standards, had been derived from acceptable values of fundamental constants, or had been derived by the ratio type of self-calibration techniques. Reference standards used in the calibration of DOD equipment and systems had to be supported by certificates, reports, or data sheets attesting to the date, accuracy, and conditions under which the results were obtained. Although other U.S. government agencies established requirements for traceability to NBS, the requirements of MIL-STD-45662A were the most significant, since DOD was one of the largest purchasers of goods and services in the world.

During the 1970s, many organizations asked NBS for advice on how to comply with traceability requirements, since these requirements usually included language to have measurements "traceable to NBS." Unfortunately, these agencies as well as industrial organizations included traceability requirements without considering the technical or scientific implications of traceability, or even having a consensus definition as to what it meant. As a result, there was confusion, misinformation, misinterpretation, and lack of agreement among the diverse and large community of testing and/or calibration facilities, vendors, suppliers, and end-users of the products or services covered by the requirements of a contract or regulation.

Since contract auditors were not metrologists, they did not go beyond checking to see whether the contractor had calibration certificates from NBS on file. A paper trail tracing back to NBS primary or reference standards or SRMs can be of dubious value in ensuring that actual measurements are valid.

A story that may well be apocryphal was widely repeated to illustrate the shortcomings of the conventional view of traceability: There was a defense contractor that had a set of gage blocks calibrated by NIST at regular intervals and used the calibration certificates to prove to auditors that the company was "traceable to NBS" for its dimensional measurements. Allegedly, the company never opened or used the box of gage blocks. Each time the box returned to NBS for recalibration the seals were found unbroken. Auditors focused only on documentation for instruments or artifacts—not the entire measurement process. A measurement device might be properly calibrated and in good working order, but if the operator did not know how to use it properly, the resulting measurements could still be inaccurate.

Also during the 1960s and 1970s, internationally-renowned NBS statisticians Churchill Eisenhart, Harry Ku, and Joe Cameron taught that to have valid measurements, one must be able to quantify the random error (Type A), as well as put bounds on the maximum possible systematic errors (Type B) of the measurements, and to show that the total uncertainty is sufficiently small to accomplish the purpose at hand. The concept of the "Measurement Assurance Program," or MAP, was developed during this era, and NBS worked with its customers to help them understand it. By the late 1970s, the clamor in both industry and government was increasing to define better what traceability meant. Within NBS, it was clear that unless measurement uncertainty is quantified and monitored, traceability is meaningless. The American Society for Testing and Materials, ASTM, was forming a new Committee E-46 on Quality Systems, and the topic of traceability was one of the items the committee was struggling with. Committee E-46 was looking to NBS to help define how traceability related to quality systems.

Brian Belanger, an electrical engineer, came to NIST in 1977 to work in the Institute for Basic Standards (IBS), then directed by Arthur McCoubrey. Shortly after Belanger arrived at NBS, Joe Cameron retired as Chief of the Office of Measurement Services, and Belanger was named to replace him. Cameron had introduced him to Eisenhart and Ku, and he quickly became an advocate for measurement assurance programs. Discussions with John Simpson, another NBS manager who was particularly articulate on the inadequacies of the conventional approach to traceability, also served to convince Belanger that things needed to change. Having read some earlier papers on traceability by Joe Cameron and Harmon Plumb [2-3], and with the encouragement of Art McCoubrey, Belanger set about to write an article on traceability for *ASTM Standardization News*.

Fig. 2. Joseph Cameron.

The paper stressed that to be meaningful, traceability needed to evolve beyond an emphasis on auditing a paper trail and, instead, focus on whether the measurement process was in control with regard to random errors and drift, and whether the measurements had an acceptably small possible offset from national standards. In other words, the thrust of the paper was to insert the idea of measurement assurance into the traceability debate. The key concepts are that measurement uncertainty must be quantified relative to national standards and that data must be produced on a continuing basis to demonstrate that the measurement process is in a state of statistical control. Belanger recommended that traceability be defined as follows:

"Traceability to designated standards (national, international, or well-characterized reference standards based upon fundamental constants of nature) is an at tribute of some measurements. Measurements have traceability to the designated standards if and only if scientifically rigorous evidence is produced on a continuing basis to show that the measurement process is producing measurement results (data) for which the total measurement uncertainty relative to national or other designated standards is quantified."

Full acceptance of the definition proposed by Belanger became an accomplished fact at NBS and, ultimately, at the NMI's of other major industrial countries. Indeed, it is a necessary and essential condition that a world class NMI produce "scientifically rigorous evidence" on a continuing basis to show that its measurement processes are quantifiable and in a state of statistical control. At the other end of the measurement hierarchy—the manufacturer's shop floor, small testing facilities, and calibration vendors—"scientifically rigorous evidence" and quantifiable uncertainty were more nearly perceived as a barrier to traceability and an unnecessary luxury. Even so, there was strong support for the concept within NBS and staff reinforced the message in their own writings and talks during the 1980s. By 1986, an American Society for Quality Control (ASQC) writing group had developed ANSI/ASQC Standard M1 on Calibration Systems, which included a definition of traceability based on the concept of quantified uncertainty.

As support for the new approach to traceability began to grow in the United States, metrologists in other countries gradually began to endorse the approach as well. In the 1993 edition of the International Vocabulary of Basic and General Terms in Metrology (VIM) traceability is defined as follows:

"Property of the result of a measurement or the value of a standard whereby it can be related to stated references, usually national or international standards, through an unbroken chain of comparisons all having stated uncertainties."

While this definition does not explicitly embrace the concept of measurement assurance as a means of demonstrating that the measurement process is in a state of statistical control, it does embrace all of the other basic elements of the approach to traceability outlined in Belanger's paper. The VIM definition has perhaps gained the widest acceptance on both the national and international scene of any definition of traceability published to date.

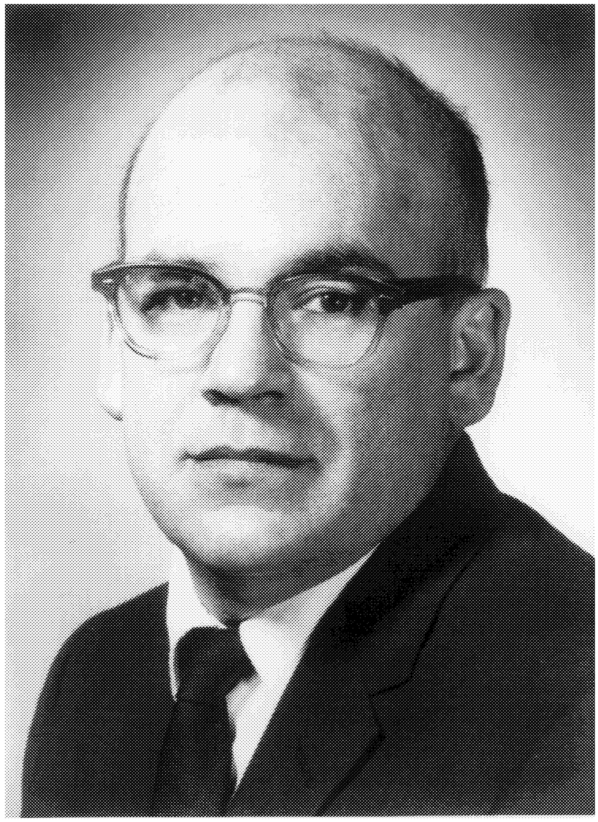

Fig. 3. Harmon Plumb.

Thinking on the subject continued to progress at NIST as well as elsewhere. In 1992 the Testing and Evaluation Committee of ASTM invited the Calibration Program to reflect on the decade of the 1980s and to give an assessment of new developments and trends in traceability. Ernest Garner and Stanley Rasberry of the Office of Measurement Services responded to the invitation with a paper [4], which was published in the ASTM Journal of Testing and Evaluation. Identifiable trends included: increased attention to defining and minimizing the number of measurement comparisons in the unbroken chain to NIST; local representation (i.e., intrinsic standards such as the Josephson Junction for voltage and the Quantum Hall Effect for resistance); greater use and acceptance of laboratory accreditation; significant movement to quality programs based upon internationally recognized documentary standards; and the emergence of multinational, regional trade organizations requiring improved mutual recognition, equivalency of measurements and confidence among accreditation and metrology bodies. The authors also called attention to the critical implications of transferring SI units from the point of realization (usually an NMI) to an end user, who then had to determine what uncertainty was required to satisfy the intended purpose.

Maintaining a measurement quality assurance system to ensure that total uncertainty was within specified limits at all times was recommended.

While the use of the concept of traceability was becoming more popular, its implementation was still less than ideal. During his years as Leader of the NIST Pressure Group in the late 1980s and early 1990s, Charles Ehrlich developed both an appreciation and frustration for the way that traceability was being treated in that field. He became increasingly concerned that while uncertainty in measurement was being recognized as an important component of traceability, the use of appropriate measurement assurance methods both before and after calibrations was not. Further, the community was taking different approaches to what constituted traceability, creating considerable confusion. In the mid-1990s Ehrlich teamed with Rasberry, who shared many of the same concerns, to move the traceability discussion ahead. Their collaboration produced the next and most recent contribution to this NBS/NIST traceability series, the paper *Metrological Timelines in Traceability* [5]. Besides appearing in the NIST Journal of Research, this paper was also published in the journal Metrologia and, by invitation, in the Bulletin of the International Organization of Legal Metrology (OIML). Evidence of the continued interest in the metrology community of furthering the traceability concept came when this paper received the "best paper" award in the management category at the 1997 National Conference of Standards Laboratories Symposium.

The *Metrological Timelines* paper, which calls attention to the time element in controlling measurements (implied in the term "continuously" in Belanger's paper), explains the modern view of traceability very thoroughly, and introduces the concept of the metrological timeline. Ehrlich and Rasberry emphasize that what is a valid "traceable" measurement today may not be valid at a later time unless the organization(s) in the unbroken chain have in place adequate measurement assurance procedures. They further emphasize that the responsibility for traceability resides not just with the NMI, but also with the lower level laboratory, which must put in place a process for guaranteeing that an instrument sent to an NMI for calibration is behaving the same way when it returns from the NMI as it did before it was sent. Ehrlich and Rasberry also elaborate on other aspects of traceability, such as multiple routes to traceability, equivalence of measurements, recalibration intervals and the use of intrinsic standards in traceability. They propose that the key questions to be answered concerning traceability are: " What corrections should be applied to a measurement result obtained at a given time with my instrument to match the result that would be obtained using the instrument (standard) to which traceability is desired? What is the uncertainty of this corrected measurement result?" They also note that "An uncertainty cannot be stated rigorously without demonstrated traceability."

Since traceability is an important ingredient of laboratory accreditation and international mutual recognition agreements, the importance of adopting a rigorous approach to traceability goes well beyond issues such as contractual compliance for defense contractors. Organizations such as NORAMET (The North American Metrology Cooperation), ISO, and OIML are using approaches that stress that quantified measurement uncertainty is a key aspect of traceability. While great progress has been made, traceability continues to evolve.

Brian C. Belanger holds a B.S. from Caltech and a Ph.D. from the University of Southern California, both in Electrical Engineering. After four years at the General Electric Research and Development Center, working in superconductivity, cryogenics, and high voltage dielectrics, he joined the Federal government in 1972, becoming a program manager at the Atomic Energy Commission (AEC), with responsibility for developing high capacity superconducting underground power cables. He joined NBS in 1977, soon becoming Chief of the Office of Measurement Services in the Institute for Basic Standards. Belanger served as NBS' liaison to the Department of Defense (early 1980s), as Associate Director for Program Development in the Center for Electronics and Electrical Engineering (1987-90), as a Senior Technology Advisor to the Assistant Secretary for Technology Policy at the Commerce Department (1990), as the Deputy Director of the Advanced Technology Program (1991-98), and finally, as Executive Director of the NIST Visiting Committee on Advanced Technology.

Joseph M. Cameron received his B.S. in Mathematics in 1942 from the University of Akron and an M.S. in Statistics in 1947 from North Carolina State. He joined the NBS staff in 1947, and served as Chief of the Statistical Engineering Section in the NBS Applied Mathematics Division. In 1969 Cameron was appointed Chief of the newly formed Office of Measurement Services in the Institute for Basic Standards. He retired in 1977 and died in January 2000.

After completing his B.S. in Physics at the University of Miami and his Ph.D. in Surface Physics at the University of Pennsylvania, and working for five years in the semiconductor industry, Charles D. Ehrlich joined the Pressure and Vacuum Group at NBS/NIST in 1984, working to help establish the first leak calibration program. He served as Group Leader of the Pressure Group from 1987-1994, followed by two years in the

NIST Program Office. Ehrlich is currently Chief of the Technical Standards Activities Program.

Ernest L. Garner has a measurement background at NIST/NBS in high accuracy mass spectrometry of the elements to include atomic weight determinations, elemental abundance by isotope dilution, natural isotopic variations of the elements, and characterizing the sources of uncertainty in thermal ionization mass spectrometry.

Harmon H. Plumb received his B.A. in Physics from Middlebury College in 1947, and his M.S. and Ph.D. degrees from Northwestern University in 1948 and 1954, respectively. He joined the NBS Cryogenic Physics Laboratory in 1955, working to establish low temperature scales. Prior to joining NBS, Plumb performed basic research on ice single crystals for the Corps of Army Engineers, and from 1949 to 1950 he was a physics instructor at Miami University, Oxford, Ohio. He died in 1985.

Stanley D. Rasberry began his career at NIST in 1959, while studying for a degree in physics at the Johns Hopkins University. He developed and applied spectrometric techniques to the chemical analysis of a wide variety of materials, especially those intended as standard reference materials (SRMs). During the late 1960s, he was privileged to work with Joe Cameron to improve and computerize procedures for quantifying the homogeneity of candidate SRMs. The collaboration provided an appreciation that SRMs could serve as links for the traceability of chemical measurements to national standards. Later, Rasberry became Chief of the Standard Reference Materials Program and subsequently Director of the Office of Measurement Services, where he also directed NIST efforts in standard reference data, weights and measures, calibration services, and the National Voluntary Laboratory Accreditation Program. The Measurement Services assignment gave him the opportunity to explore and develop traceability concepts together with Garner and Ehrlich. Following retirement, in 1997, Rasberry has continued to consult with NIST and has explored new modes of demonstrating traceability, including the use of proficiency testing with results being compared to those of an NMI.

Prepared by Brian Belanger, Stanley Rasberry, Ernest Garner, Carroll Brickencamp, and Charles Ehrlich.

Bibliography

[1] Brian C. Belanger, Traceability: An Evolving Concept, *ASTM Stand. News* **8** (1), 22-28 (1980).

[2] J. M. Cameron and H. Plumb, *Traceability—with Special Reference to Temperature Measurement, Society of Automotive Engineers,* Report 690428, National Air Transportation Meeting (1969).

[3] Joseph M. Cameron, Traceability? *J. Qual. Technol.* **7**, 193-195 (1975).

[4] Ernest L. Garner and Stanley D. Rasberry, What's New in Traceability, *J. Test. Eval.* **21**, 505-509 (1993).

[5] Charles D. Ehrlich and Stanley D. Rasberry, Metrological Timelines in Traceability, *J. Res Natl. Inst. Stand. Technol.* **103**, 93-105 (1998).