

# *Realistic Evaluation of the Precision and Accuracy of Instrument Calibration Systems*

Prior to the publication of *Realistic Evaluation of the Precision and Accuracy of Instrument Calibration Systems* by Churchill Eisenhart [1], the terms “precision” and “accuracy” were used in a qualitative manner to characterize measurements. These terms appeared in many American Society for Testing Materials (ASTM) standards long before any common agreement or understanding had been reached as to their meanings and consequences. Circa 1950, individuals and organizations began concerted efforts to right this situation. Churchill Eisenhart was drawn to this issue as it related to calibrations, which he called refined measurement methods. As Chief of the Statistical Engineering Laboratory (SEL), Applied Mathematics Division, he set out to put the concepts of accuracy and precision on a solid statistical basis for NBS scientists and metrologists.

His paper on the subject, published in 1961 [1], was to become the preeminent publication on the subject. With impeccable scholarship and commitment to detail, Eisenhart synthesized his own work [2] and the writings of statistical theorists and practitioners, Walter Shewhart [3], Edwards Deming, Raymond Birge [4], and R. B. Murphy [5], into concepts of quality control that could be applied to measurement processes.

Three basic concepts in the paper were immediately accepted by metrologists at NBS, namely: (1) a measurement process requires statistical control; (2) statistical control implies control of both reproducibility and repeatability; and (3) a measurement result requires an associated statement of uncertainty that includes any possible source of bias.

In this paper, for the first time, measurements themselves were described as a process whose output can be controlled using statistical techniques. Eisenhart reinforced the conclusion, probably first drawn by Murphy [5], that “Incapability of control implies that the results of measurement are not to be trusted as an indication of the physical property at hand—in short, we are not in any verifiable sense measuring anything”—when he says, “a measurement operation must have attained what is known in industrial quality control language as a state of statistical control . . . before it can be regarded in any logical sense as measuring anything at all.”

Eisenhart’s paper, coupled with work by other SEL statisticians, had a lasting and profound effect on

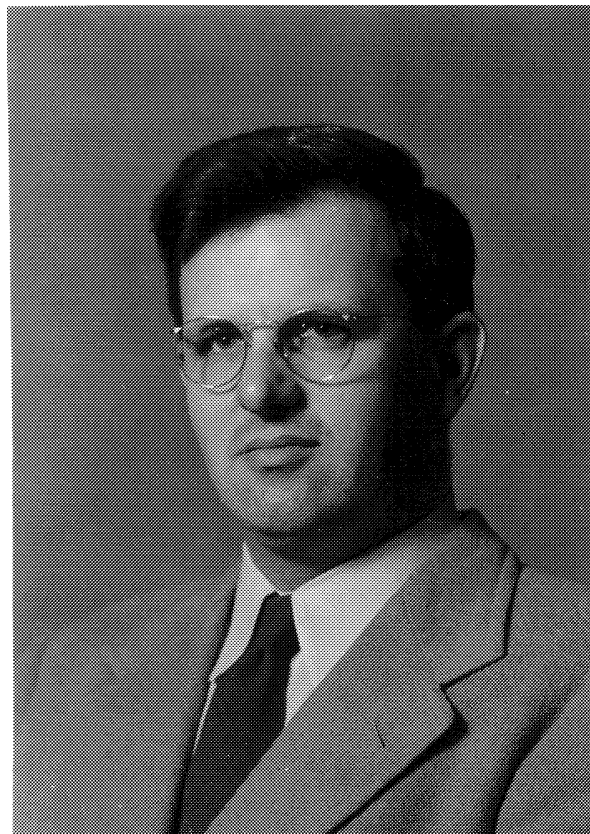


Fig. 1. Churchill Eisenhart.

measurement processes at NBS/NIST and throughout the metrology community. W. J. Youden revolutionized interlaboratory testing with methods for ruggedness testing [6] and for quantifying bias in test methods [7] and scientific measurements [8]. In his work with industrial chemists and ASTM committees, Youden left a huge body of literature on the subject of bias. His papers, which are too numerous to cite, have a common theme in the use of experimental design to shed light on sources of error in a measurement process. He was especially interested in interlaboratory testing as a means of uncovering biases in measurement processes [9], and the so-called Youden plot [10] has become an accepted design and analysis technique throughout the world for interlaboratory comparisons.

Statistical activity at NBS in the 1950s was characterized by the development of experimental designs; in the

late 1950s, with the advent of electronic computing, Joseph Cameron created calibration designs with provisions for check standards for the NBS calibration laboratories. Cameron, Youden, and Eisenhart then merged the check standard concept with quality control procedures in Eisenhart's paper to form a cohesive practice, known as measurement assurance [11-13], as a means of tying measurement results to a reference base and quantifying the uncertainty relative to the reference base. The first documentation of a measurement assurance program in a NBS calibration laboratory appears to be a tutorial by Paul Pontius and Joseph Cameron [14] on mass calibrations. Measurement assurance programs now abound in metrology and are regularly applied to measurements as diverse as dimensional measurements of gage blocks standards and semiconductor devices [15-16].

Eisenhart's exposition of sources of error in a measurement process led to the accepted practice of the day for reporting uncertainty as described by Harry Ku [17], and a paper co-authored with Ron Collé and Ku [18] was a forerunner of the 1993 *ISO Guide to the Expression of Uncertainty in Measurement* [19] and the companion NIST guideline by Barry Taylor and Chris Kuyatt [20].

Churchill Eisenhart was brought to NBS from the University of Wisconsin in 1946 by Edward Condon, Director of NBS, to establish a statistical consulting group to "substitute sound mathematical analysis for

costly experimentation." He was allowed to recruit his own staff and, over the years, he brought many notable and accomplished statisticians to SEL. He served as its Chief from 1947 until his appointment as Senior Research Fellow in 1963. He retired in 1983, and his final contribution to NIST was the formation of the Standards Alumni Association, which he headed until his death in 1994.

In its early days, SEL was drawn into outside studies as NBS became more involved in industrial activities. The study that brought the most controversy to NBS and the most recognition to Eisenhart's group was the AD-X2 battery additive case. The NBS Director, A.V. Astin, had been pressured by various senators and the battery additive producer to run a test of the additive. Under extreme time constraints, the statisticians came up with appropriate experimental designs for the tests [21] and assigned the treatment of additive or no-additive to 32 batteries blindly and at random. The experiments, run by the Electricity Division, confirmed that the additive had no significant positive effect on batteries, but in what was quickly to become an interesting sidelight of history, the Assistant Secretary of Commerce for Domestic Affairs announced that Astin had not considered the "play of the marketplace" in his judgment and relieved him as Director of NBS. Eventually, the National Academy of Sciences was called in to review NBS's work, which was labeled first rate, and Astin was reinstated [22].

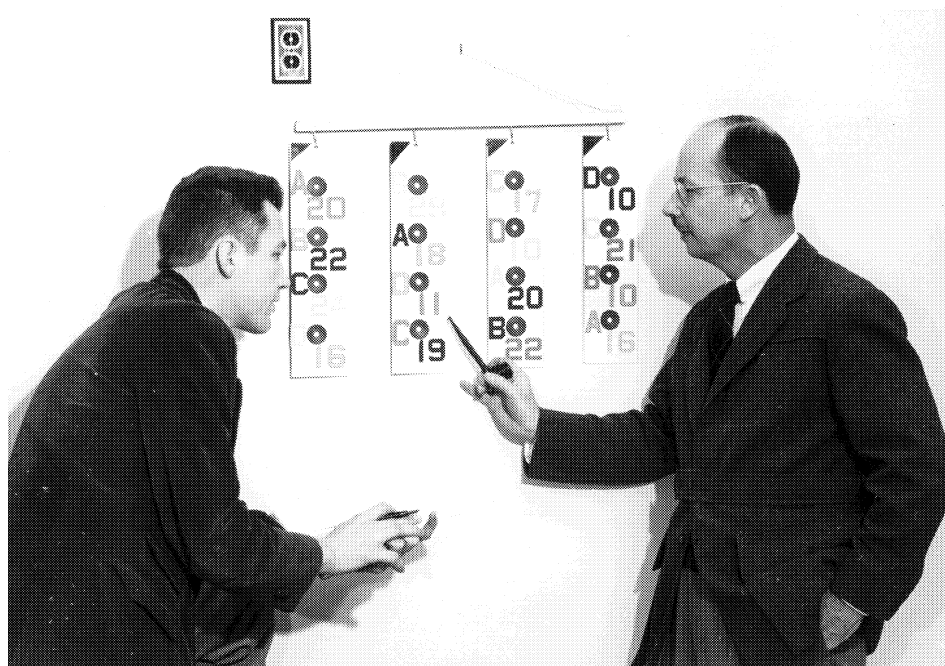


Fig. 2. Joseph Cameron and Jack Youden of the Statistical Engineering Laboratory explaining a measurement design.

Over his long and illustrious career, Eisenhart was awarded the U. S. Department of Commerce Exceptional Service Award in 1957; the Rockefeller Public Service Award in 1958; and the Wildhack Award of the National Conference of Standards Laboratories in 1982. He was elected President of the American Statistical Association (ASA) in 1971 and received the Association's Wilks Memorial Medal in 1977. Eisenhart was honored with an Outstanding Achievements Award of the Princeton University Class of 1934 and with Fellowships in the ASA, the American Association for the Advancement of Science, and the Institute of Mathematical Sciences. He was a long-time member of the Cosmos Club.

In this later years, Eisenhart indulged his interest in the history of statistics, and particularly in the evolution of least-squares. He corresponded regularly with those who had like interests. In a memorial lecture given in his honor at the National Institute of Standards and Technology on May 5, 1995, Stephen Stigler [23] says that "I wrote to him that he had set the standard for scholarly research in our field, and that is how I thought of him—the standard."

*Prepared by M. Carroll Croarkin.*

## Bibliography

- [1] Churchill Eisenhart, Realistic Evaluation of the Precision and Accuracy of Instrument Calibration Systems, *J. Res. Natl. Bur. Stand.* **67C**, 161-187 (1963).
- [2] C. Eisenhart, The Reliability of Measured Values—Part I, Fundamental Concepts, *Photogramm. Eng.* **XVIII**, 542-554 (1952).
- [3] Walter A. Shewhart, *Statistical Method from the Viewpoint of Quality Control*, The Graduate School, U.S. Department of Agriculture, Washington, DC (1939).
- [4] W. Edwards Deming and Raymond T. Birge, On the Statistical Theory of Errors, *Rev. Mod. Phys.* **6**, 119-161 (1934); with additional notes dated 1937 (The Graduate School, Department of Agriculture, Washington, DC).
- [5] R. B. Murphy, On the Meaning of Precision and Accuracy, *Mater. Res. Stand.* **1**, 264-267 (1961).
- [6] W. J. Youden, The Collaborative Test, presented at the referees' Seventy-sixth Annual Meeting of the Association of Official Agricultural Chemists, Washington, DC, Oct. 16, 1962.
- [7] W. J. Youden, How to Evaluate Accuracy, *Mater. Res. Stand.* **1**, 268-271 (1961).
- [8] W. J. Youden, Systematic Errors in Physical Constants, *Phys. Today* **14** (9), 32-43 (1961).
- [9] W. J. Youden, Experimental Design and ASTM Committees, *Mater. Res. Stand.* **1**, 862-867 (1961).
- [10] W. J. Youden, The Sample, the Procedure and the Laboratory, *Anal. Chem.* **32** (13), 23A-37A (1960).
- [11] J. M. Cameron, *Measurement Assurance*, NBS Internal Report 77-1240, National Bureau of Standards, Washington, DC (1977).
- [12] Brian Belanger, *Measurement Assurance Programs Part I: General Introduction*, NBS Special Publication 676-I, National Bureau of Standards, Washington, DC (1984).
- [13] Carroll Croarkin, *Measurement Assurance Programs Part II: Development and Implementation*, NBS Special Publication 676-II, National Bureau of Standards, Washington, DC (1984).
- [14] P. E. Pontius and J. M. Cameron, *Realistic Uncertainties and the Mass Measurement Process: An Illustrated Review*, NBS Monograph 103, National Bureau of Standards, Washington, DC (1967).
- [15] Carroll Croarkin, John Beers, and Clyde Tucker, *Measurement Assurance for Gage Blocks*, NBS Monograph 163, National Bureau of Standards, Washington, DC (1979).
- [16] Carroll Croarkin and Ruth N. Varner, *Measurement Assurance for Dimensional Measurements on Integrated-Circuit Photomasks*, NBS Technical Note 1164, National Bureau of Standards, Washington, DC (1982).
- [17] Harry H. Ku, Expressions of Imprecision, Systematic Error, and Uncertainty Associated with a Reported Value, *Measurements & Data*, **2** (4), 72-77 (1968).
- [18] Churchill Eisenhart, Harry H. Ku, and R. Collé, *Expression of the Uncertainties of Final Measurement Results: Reprints*, NBS Special Publication 644, National Bureau of Standards, Washington, DC (1983).
- [19] *Guide to the Expression of Uncertainty in Measurement*, International Organization for Standardization, Geneva, Switzerland (1993).
- [20] Barry N. Taylor and Chris E. Kuyatt, *Guidelines for Evaluating and Expressing the Uncertainty of NIST Measurement Results*, NIST Technical Note 1297, National Institute of Standards and Technology, Gaithersburg, MD (1994).
- [21] *Battery AD-X2*, Hearings before the Select Committee on Small Business, United States Senate, Eighty-Third Congress, First Session, On Investigation of Battery Additive AD-X2, Washington, 1953.
- [22] *NBS Report 2447 on Battery Additive AD-X2*, U.S. Government Printing Office, Washington, DC (1953).
- [23] S. M. Stigler, Statistics and the Question of Standards, *J. Res. Natl. Inst. Stand. Technol.* **101**, 779-789 (1996).