Int J Thermophys (2009) 30:351–359
DOI 10.1007/s10765-008-0436-7

Proficiency Testing for Achieving Accreditation in Thermometry

S. Bruce · G. F. Strouse

Published online: 5 June 2008

Abstract The establishment of traceability is an essential component of laboratory accreditation for calibration laboratories accredited by the National Voluntary Laboratory Accreditation Program (NVLAP). Traceability requires that an unbroken chain of comparisons to national standards with stated and documented uncertainties be established. The use of proficiency testing is an important tool for validating uncertainty claims contained within a Scope of Accreditation. As such, the National Institute of Standards and Technology (NIST) Thermometry Group conducts proficiency testing for NVLAP-accredited contact thermometry laboratories. These proficiency tests include thermometric fixed points, standard platinum resistance thermometers, industrial platinum resistance thermometers, thermistors, thermocouples, and liquid-in-glass thermometers. The NIST Thermometry Group proficiency tests scale inversely in difficulty as a function of the uncertainties contained within a facility's Scope of Accreditation. The offered proficiency tests are designed to establish confidence in the participant's calibration capabilities and uncertainty claims without being an undue burden to the calibration facility. The NIST Thermometry Group proficiency tests involve measurements of NIST-owned artifact(s) at NIST, measurements of the artifact(s) at the participant's calibration facility, and then a second set of measurements at NIST. On completion of the measurement phase of a proficiency test, a detailed report is generated to give the difference in temperature realization between that of the participant and NIST and the degree of equivalence. As part of the goal to improve

S. Bruce (✉)
National Voluntary Laboratory Accreditation Program, National Institute of Standards and Technology, 100 Bureau Drive, MS 2140, Gaithersburg, MD 20899-2140, USA
e-mail: sbruce@nist.gov

G. F. Strouse
Thermometry Group, National Institute of Standards and Technology, 100 Bureau Drive, MS 2140, Gaithersburg, MD 20899-2140, USA

the participant's measurement capabilities, the report also gives suggested solutions to any identified measurement issues that need attention. The role of proficiency testing as part of the process in achieving NVLAP accreditation in thermometry is discussed in the article. Additionally, the descriptions of the NIST Thermometry Group unbiased third-party proficiency tests and some unidentified participant results are given.

Keywords Accreditation · Degree of equivalence · ITS-90 · National Voluntary Laboratory Accreditation Program · NVLAP · Proficiency test · Scope of Accreditation/Temperature · Thermometry · Traceability

1 Introduction

The National Voluntary Laboratory Accreditation Program (NVLAP) is a voluntary, nonprofit, fee-supported program created to accredit laboratories that are found competent to perform specific tests or calibrations within stated uncertainties. The program was established by the National Institute of Standards and Technology (NIST) in 1976 to serve the needs of the government and private sector (industry, consumers, and other stakeholders) by fostering and promoting a uniformly acceptable base of professional and technical competence in the laboratory community, and to facilitate and promote acceptance of calibration and test results between countries to avoid barriers to trade. The combination of testing and calibration laboratory accreditations by NVLAP provides an infrastructure of competent measurement laboratories supporting domestic and international trade and conformity assessment activities. The program provides an unbiased third-party evaluation and recognition of performance, as well as expert technical guidance to upgrade laboratory performance. NVLAP procedures are contained in Part 285 of Title 15 of the U.S. Code of Federal Regulations [1].

Laboratory accreditation is further supported through the International Committee of Weights and Measures (CIPM) mutual recognition arrangement (MRA) among national metrology institutes (NMIs), to which NIST is a signatory and an active participant [2]. Additionally, NVLAP has entered into MRAs with other laboratory accreditation bodies through the Asia Pacific Laboratory Accreditation Cooperation (APLAC), and the International Laboratory Accreditation Cooperation (ILAC). These arrangements provide NVLAP customers with global acceptance of their calibration and test data. Laboratory accreditation MRAs and the CIPM MRA are complementary agreements that contribute to the reduction of trade barriers. The CIPM MRA ensures that signatory NMIs disseminate measurement results that are traceable to the International System of Units (SI), while laboratory accreditation assures that accredited laboratories are capable and competent to provide measurement results traceable to the SI through standards maintained by those NMIs. Hence, the concept 'one test, one calibration, sell anywhere' can be realized.

The NVLAP Calibration Laboratory Accreditation Program (LAP) [3], established in 1992 at the request of the National Conference of Standard Laboratories (NCSL), now known as NCSLI, supplies accreditation in the following metrology fields:

- Dimensional
- Electromagnetics—Direct current/Low frequency

- Electromagnetics—Radio frequency/Microwave
- Ionizing radiation
- Mechanical
- Optical radiation
- Thermodynamics
- Time and frequency

Program-specific technical requirements for achieving NVLAP accreditation in each of the listed metrology fields is found in the NVLAP 150 series handbooks [4].

NVLAP currently has 30 calibration laboratories that are accredited for thermodynamic parameters. What makes NVLAP distinct amongst the large number of accreditation bodies throughout the United States is our direct relationship with NIST metrologists and their ability to participate in the accreditation process as technical experts and in proficiency testing (e.g., NIST Thermometry Group—Gregory Strouse, Dean Ripple, and Karen Garrity). The ability of NVLAP to function efficiently and effectively in partnership with other organizations within NIST is key to the value-added services that we offer to our customers. As of 2002, the NIST Thermometry Group has provided NVLAP with technical assistance, third-party proficiency tests (PTs) and artifacts, and qualified technical assessors.

The establishment of traceability is one essential component of laboratory accreditation for calibration laboratories accredited by NVLAP. Traceability requires that an unbroken chain of comparisons to national standards with stated and documented uncertainties be established. NIST assures the traceability of results of measurements or values of standards that NIST itself provides, either directly or through an official NIST program or collaboration. Other organizations are responsible for establishing the traceability of their own results or values to those of NIST or other stated references. NIST has adopted this policy statement to document the NIST role with respect to traceability [5]. Additional components of accreditation include a quality system compliant with the International Organization for Standardization and the International Electrotechnical Commission, ISO/IEC 17025:2005, periodic third party technical and quality system assessment, correction of nonconformities encountered during an assessment, and successful completion of a proficiency test. This article highlights three of the six types of NVLAP thermometry proficiency tests that are offered: thermocouples, industrial platinum resistance thermometers (IPRTs), and standard platinum resistance thermometers (SPRTs).

2 Proficiency Testing

Proficiency testing is an important tool for validating uncertainty claims contained within a Scope of Accreditation. Figure 1 shows the nominal steps followed during a NVLAP-administered proficiency test. NVLAP does not impose specific calibration procedures or lower limits on the laboratory's submitted Scope of Accreditation and uncertainty budget. NVLAP allows for all scientifically valid calibration schemes and requires that a laboratory derive and document its uncertainties. It is the laboratory's responsibility to properly validate those uncertainty claims through a proficiency test.

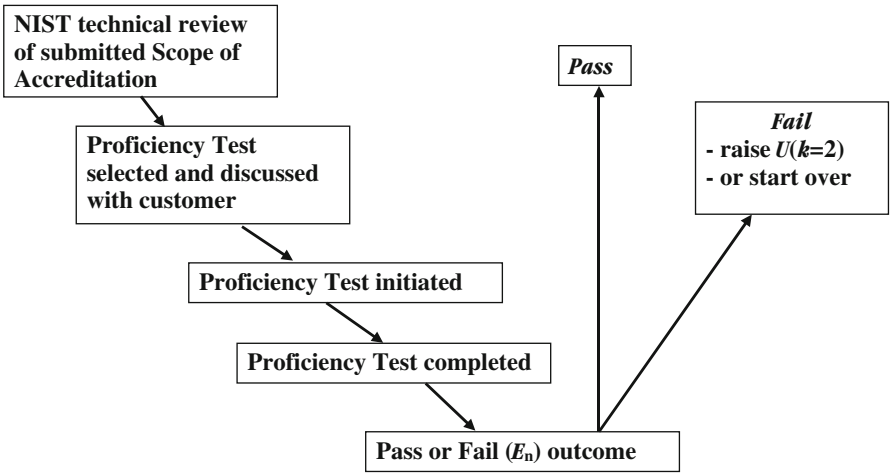

Fig. 1 Nominal steps of a NVLAP-administered proficiency test

In the NVLAP Calibration Laboratory Accreditation Program, proficiency testing is used as a decision point on whether or not to accredit a laboratory.

When a laboratory submits a Scope of Accreditation as part of the accreditation request, the Scope of Accreditation is technically reviewed by the pertinent NIST metrology staff. From this review, a suitable proficiency test is selected, discussed, and agreed upon between both the NVLAP program manager and the laboratory's contact. The NIST Thermometry Group proficiency tests involve measurements of NIST-owned artifact(s) at NIST, measurements of the artifact(s) at the participant's calibration facility, and in most cases a second set of measurements at NIST (see Sect. 2.1 for the exception). The proficiency test artifacts are measured at only one laboratory between NIST measurements of the artifacts, thus reducing the transfer artifact uncertainty. The NIST Thermometry Group proficiency tests are designed to scale inversely in difficulty as a function of the uncertainties contained within a facility's Scope of Accreditation. The proficiency tests are designed to establish confidence in the participant's calibration capabilities and uncertainty claims.

Table 1 lists the available NVLAP International Temperature Scale of 1990 (ITS-90) based proficiency tests provided by the NIST Thermometry Group. Within each thermometer calibration area and temperature range, various proficiency test options exist to match the requirements of the laboratory. For instance, a proficiency test for SPRTs can be modified to cover a smaller ITS-90 temperature subrange of 0 °C to 30 °C.

On completion of the measurement phase of a proficiency test, a detailed NIST Thermometry Group report is delivered to NVLAP. The report gives detailed information about the analysis of the participant's results, uncertainty budget, and the E_n value (degree of equivalence). For each test point, an E_n value is calculated by

$$E_n = \Delta T_{90}\,(\text{Participant} - \text{NIST})/U_{\text{C}}\,(k = 2), \tag{1}$$

Table 1 NVLAP available proficiency tests for contact thermometry. The NIST Thermometry Group is responsible for administering the proficiency tests to the laboratory

ITS-90 calibration area	Proficiency test artifact	Maximum temperature range (°C)	Comments
ITS-90 fixed-point cells	Laboratory's own	−190 to 962	Direct comparison with NIST
Standard Platinum Resistance Thermometers (SPRTs)	SPRT	−190 to 661	Other ranges available
Industrial Platinum Resistance Thermometers (IPRTs)	SPRT	−190 to 550	SPRT used as the transfer artifact
Thermistors	Thermistor	−50 to 90	
Thermocouples	Thermocouple	−190 to 1,100	Type K, Destructive test to artifact
Liquid-in-glass	Liquid-in-glass	−20 to 400	

where U_C is the combined expanded uncertainty of the participant, NIST, and artifact(s) [6]. An $|E_n|$ value of less than one signifies compliance for the participant. As part of the goal to improve the participant's measurement capabilities, the report also gives suggested solutions to any identified measurement issues that need attention.

2.1 Thermocouple Proficiency Test

The thermocouple proficiency test is available for Type K thermocouples to cover the temperature range of −196 °C to 1,100 °C as a representative test for all base-metal thermocouple calibrations. The NIST Thermometry Group purchased and characterized nickel/aluminum/manganese (KN) and nickel/chromium (KP) spools of 1.63 mm diameter bare wire. Inhomogeneity determination of the wire lot was included in the characterization measurements. The large wire gauge was chosen because of the increased difficulty in handling a large wire diameter and the inherent thermoelectric drift that occurs with small gauge wire when used above 800 °C. The thermocouple proficiency test is configured as a one-way Exploding Star; with each participant receiving two (2) pairs of untested Type K thermocouple wire from the NIST-characterized spools. Because this test is destructive to the wire, the participant's measurement results of the supplied wire are compared with that of the NIST characterization of the wire.

The NIST Thermometry Group characterization of the wire included the evaluation of the thermoelectric inhomogeneity and the average emf versus temperature response of the wire. The wire was tested in three different apparatuses. For the five tested thermocouples, there were no statistically significant differences in the standard deviations of the inhomgeneity results for the three apparatuses. The NIST Type A inhomogeneity and repeatability uncertainty varied from 0.02 °C at 100 °C, 0.05 °C at 400 °C, and 0.03 °C at 1,100 °C. No trend was detectable in the emf versus temperature response as a function of wire location on the spool; thus, the emf versus temperature response of any one wire sample is equivalent to the average of the tested wire samples. The thermoelectric changes (drift) ranged from 0.0 °C at 100 °C to 0.4 °C at 1,100 °C. The NIST $U(k = 2)$ ranged from 0.05 °C at 100 °C to 1.0 °C at 1,100 °C.

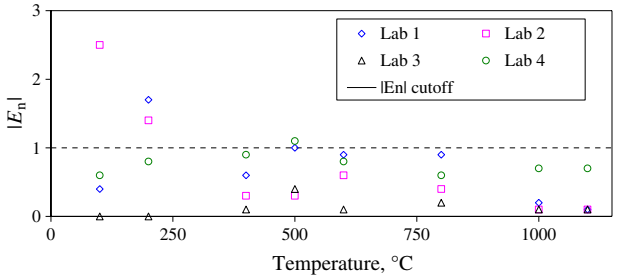

Fig. 2 Results of the NVLAP Type K thermocouple proficiency test. An $|E_n|$ value less than one signifies compliance for the participant

Figure 2 shows the Type K thermocouple proficiency test results of four participants. In particular, three of the four participants failed the proficiency test (Lab 1, Lab 2, and Lab 4) with $|E_n|$ values greater than one. On close inspection of the initial uncertainty claims for those three participants, the lack of harmonization of their uncertainty budget, primarily the inclusion of missing uncertainty components of inhomogeneity and drift of the artifact, was the major contributor to failing the proficiency test. As a consequence, the NVLAP accreditation process resulted in the adjustment of the Scope of Accreditation of the three participants to reflect the changes in their uncertainty budget (e.g., inclusion of missing components of inhomogeneity and drift of artifact) to achieve compliance.

2.2 Standard Platinum Resistance Thermometer Proficiency Test

NIST thermometry proficiency tests for SPRTs are offered over the ranges of Ar TP to Zn FP (−190 °C to 420 °C); Ar TP to Al FP (−190 °C to 661 °C); and other ranges are available as needed. The participant's Scope of Accreditation defines the test temperature range. There is a three-tier system based on uncertainty claims that allows for different levels of uncertainties. The uncertainty cutoff values are statistically derived from NMI uncertainty claims from Regional Metrology Organizations of the Asia Pacific Metrology Program (APMP), Euro-Asian Cooperation of State Metrology Institutions (COOMET), European Collaboration in Measurement Standards (EUROMET), Southern African Development Community—Cooperation in Measurement Traceability (SADCMET), and Sistema Interamericano de Metrologia (SIM). Detailed information concerning the SPRT proficiency test and accreditation requirements is found in [7].

When laboratories are seeking or maintaining NVLAP accreditation for their ITS-90 SPRT calibration services with expanded uncertainties ($k = 2$) less than or equal to the values given in Table 2 (tier 1 level), they are required to undergo a proficiency test based on the NIST Measurement Assurance Program (MAP) and the direct comparison of fixed-point cells. The MAP is a specialized proficiency test comparable in difficulty to that of an NMI Key Comparison. The MAP PT is a bilateral-type comparison that utilizes three NIST-owned SPRTs as transfer artifacts.

Table 2 Tier 1 uncertainty cutoff values ($k = 2$)

Fixed point	Uncertainty cutoff (mK)	Fixed point	Uncertainty cutoff (mK)
LN_2	1.5	In	1.2
Ar	1.5	Sn	1.0
Hg	0.6	Zn	2.0
H_2O	0.3	Al	3.4
Ga	0.5	Ag	10.0

For those fixed-point cells that require a direct comparison with the NIST reference standards, a subset will be selected by the appropriate NIST Thermometry Group technical staff member. The direct comparison of fixed point-cells methods are described in detail in [8].

The MAP PT and direct comparison of fixed-point cells are performed during alternate years. The MAP PT is performed once every three years and the direct comparison of fixed-point cells once every three years. A single SPRT proficiency test covering the temperature range of the accreditation is performed in the intervening years. Tier 1 represents the measurement capability of the top 25% of the NMIs. Gregory Strouse or Dean Ripple serves as the technical on-site assessor. As previously stated, these two assessors alternate visits to maintain a fresh perspective.

For laboratories seeking or maintaining NVLAP accreditation within the range given in Table 3 (tier 2 level), a MAP PT is required. The MAP PT is performed once every 3 years with a single SPRT proficiency test that covers the temperature range of the accreditation in the intervening years. Tier 2 represents the 25% to 75% range measurement capability of the NMIs. Gregory Strouse or Dean Ripple serves as the technical on-site assessor.

The analysis of tier 1 and tier 2 participant's results include checking the following calculations: temperature corrections for fixed-point cells, derivation of ITS-90 deviation function coefficients, derivation of thermometer calibration tables, and error at redundant fixed points (if measured). SPRT handling is also analyzed, including the stability of the SPRTs at the triple point of water (TPW). Clarity and correctness of calibration reports are checked. A comparison of $W(T_{90})$ (the resistance ratio $W(T_{90})$ is defined as $R(T_{90})/R(273.16\,\mathrm{K})$, where R is either resistance or a resistance ratio) values at each fixed point is made to verify the calculation of $|E_n|$ and to determine the impact over the ITS-90 temperature subranges. Tier 1 and tier 2 laboratories are required to complete their proficiency test prior to the on-site technical assessment.

Table 3 Tier 2 uncertainty cutoff values ($k = 2$)

Fixed point	Uncertainty cutoff (mK)	Fixed point	Uncertainty cutoff (mK)
LN_2	>1.5 to 2.3	In FP	>1.2 to 1.8
Ar TP	>1.5 to 2.3	Sn FP	>1.0 to 2.0
Hg TP	>0.6 to 0.8	Zn FP	>2.0 to 3.4
H_2O TP	>0.3 to 0.5	Al FP	>3.4 to 8.5
Ga MP	>0.5 to 0.6	Ag FP	>10.0 to 15.0

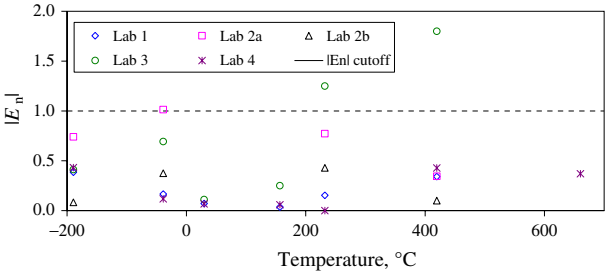

Fig. 3 SPRT proficiency test results for a representative sample of tier 1 and tier 2 laboratories. An $|E_n|$ value less than one signifies compliance for the participant

Additionally, participant laboratories providing ITS-90 SPRT calibration services are advised to have the information given in Appendix B of [7] available in preparation for their on-site assessment.

Prior to sending the PT SPRTs to the participating laboratory, NIST calibrates the SPRTs using ITS-90 fixed points. Upon receipt of the SPRTs, the laboratory then has 8 weeks to complete their measurements, calculations, and calibration reports. Finally, NIST re-measures the SPRTs on return of the artifacts. This closure of the measurement loop is critical to assessing uncertainties. It provides an assessment of artifact drift; the SPRT should repeat to within NIST measurement uncertainties. If not, then an additional transfer artifact drift is assessed and included as an uncertainty or cause for repeat of the proficiency test.

Figure 3 shows the MAP PT results for a representative sample of tier 1 and tier 2, laboratories (two tier 1 and two tier 2). For each participant, the results are shown as the average of the three SPRT results. One participant (Lab 2a) failed its initial proficiency test. For that participant, harmonization of the uncertainty budget and adjustment of the measurement techniques were made before another proficiency test was performed. As a result, the participant improved its measurement procedures and the Scope of Accreditation was adjusted by NVLAP to reflect the changes in the uncertainty budget. The implemented changes resulted in the participant passing its second proficiency test (Lab 2b). One participant (Lab 3) failed the initial proficiency test, and its Scope of Accreditation is on hold until successful completion of a MAP PT.

Tier 3 is for those laboratories that are seeking or maintaining NVLAP accreditation for their ITS-90 SPRT calibration services with expanded uncertainties ($k = 2$) greater than the values given in Table 3. A tier 3 proficiency test uses a single SPRT to cover the temperature range of the accreditation. The full temperature range of accreditation is tested once every three years. Normally, a non-NIST NVLAP technical assessor is used for the on-site assessment.

Figure 4 shows the proficiency test results for a representative sample of tier 3 laboratories. One participant (Lab 2) failed its proficiency test at one fixed point. To achieve compliance, the participant adjusted its uncertainty budget and NVLAP subsequently approved their Scope of Accreditation.

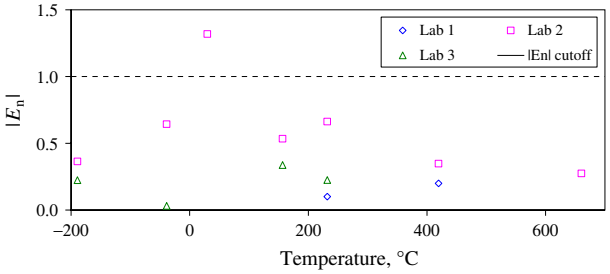

Fig. 4 SPRT proficiency test results for a representative sample of tier 3 laboratories. An $|E_n|$ value less than one signifies compliance for the participant

3 Conclusions

Outcomes of proficiency tests determine any mathematical issues in participant's calculations, identify and correct any measurement issues that may exist, and confirm uncertainty claims by a calibration facility. The proficiency testing described here also provides measurement traceability. Proficiency testing done well must adequately test the laboratory's Scope of Accreditation. This can be a sampling of the submitted Scope of Accreditation, but the proficiency test must be capable of sufficiently testing the participant lab's uncertainty claims. Proficiency testing helps to assess the laboratory's technical competency. The following attributes of the laboratory's quality management system are addressed in proficiency testing activities: traceability, equipment, uncertainty claims (verification of the Scope of Accreditation), technical staff capabilities, calibration procedures and test methods, data analysis, and quality assurance. When proficiency testing is completed prior to the on-site assessment, the assessor can perform the assessment knowing a great deal about the laboratory's capabilities. Satisfactorily passing a proficiency test is one data point in the accreditation decision process. Throughout this process, a number of laboratories have had to adjust their uncertainty claims.

References

1. http://www.gpoaccess.gov/cfr/index.html
2. "Mutual recognition of national measurement standards and of calibration and measurement certificates issued by national metrology institutes," CIPM, Paris (1999)
3. http://ts.nist.gov/Standards/upload/What-is-the-NVLAP.pdf
4. http://ts.nist.gov/Standards/Accreditation/handbook.cfm
5. http://ts.nis.gov/Traceability/nist_traceability_policy-external.cfm
6. ISO/IEC Guide (1997) 43–1
7. NVLAP Lab Bulletin, LB-10-2004 (Rev. 2007), http://ts.nist.gov/Standards/Accreditation/upload/LB_10_04_rev_2007_02_21.pdf
8. G.F. Strouse, in *Proceedings of TEMPMEKO 2004, 9th International Symposium on Temperature and Thermal Measurements in Industry and Science*, ed. by D. Zvizdić, L.G. Bermanec, T. Veliki, T. Stašić (FSB/LPM, Zagreb, Croatia, 2004), pp. 879–884