The Uncertain State of Uncertainty in Industry

Dilip A. Shah
President, E = mc^3 Solutions
Excellence = **M**easurement, **C**alibration, **C**omputers, **C**onsulting **S**olutions
197 Great Oaks Trail #130
Wadsworth, Ohio 44281-8215
Tel: 330-328-4400 Fax: 330-336-3974
E-mail: emc3solu@aol.com

1.0 INTRODUCTION

The ISO 17025 laboratory accreditation standard requires measurement uncertainty estimates of calibration and test processes for the laboratory's scope of accreditation. Measurement uncertainty budget estimates are examined by accrediting body assessors before accreditation is granted to a laboratory. Because of the requirement by the international standard, the laboratories are compelled/forced/dragged to understand the process of estimating measurement uncertainty as outlined in the Guide to Uncertainty of Measurement (herein referred to as GUM).

It is important to note the reliance of industry and the world commerce on good measurements. Without good measurements, the everyday commercial activities that we take for granted would come to its knees and chaos would develop. Yet, industry in general takes good measurement practices for granted and is ignorant of measurement uncertainty associated with making a measurement. Improving technology and innovation continues the drive towards better measurement instrumentation at less cost than the previous versions. Adding to the confusion are old vestiges of 10:1 or more recently, 4:1 Test Accuracy Ratios (herein referred to as TAR) requirements spelled out in calibration contracts, old standards and still practiced/required by the industry.

Too often the end-user of the equipment focuses on getting the measurement equipment calibrated for another year and having a sticker on the equipment attesting to that fact. This narrow focus misses out on an important aspect of the measurement process, the uncertainty associated with making a measurement and applying it correctly.

This paper will identify the GUM requirements for estimation of Measurement Uncertainty and how the requirements are misstated or wrongly interpreted by the industry. The paper also discusses the myths surrounding the Test Accuracy Ratios in today's calibration environment and why TARs should be relegated to its place in history.

2.0 GENERAL DEFINITIONS

The Vocabulary of International terms in Metrology (Herein referred to as VIM) defines the following:

- **Accuracy** - The closeness of agreement between a test result and the accepted reference value.
- **Bias** - The difference between the expectation of the test results and an accepted reference value.
- **Drift** is a slow change of a metrological characteristic of a measuring instrument.
- **Error** (of indication) of a measuring instrument is the indication of a measuring instrument minus a 'true' value of the corresponding input quantity, i.e. it has a sign
- **Laboratory bias** - the difference between the expectation of the test results from a particular laboratory and an accepted reference value.
- **Measurand** - particular quantity subject to measurement.
- **Measurement** - set of operations having the object of determining a value of a quantity.
- **Metrology** - the science of measurement.
- **National (measurement) standard** - standard recognized by a national decision to serve, in a country, as the basis for assigning values to other standards of the quantity concerned.
- **Precision** The closeness of agreement between independent test results obtained under stipulated conditions.

- **Resolution** - smallest change of measured quantity which changes the indication of an measuring instrument
- **Repeatability** - precision under repeatability conditions
- **Repeatability conditions** - where independent test results are obtained with the same method on identical test items in the same laboratory by the same operator using the same equipment within short intervals of time.
- **Reproducibility** - precision under reproducibility conditions
- **Reproducibility conditions** - where test results are obtained with the same method on identical test items in different laboratories with different operators using different equipment
- **Stability** refers to the ability of a measuring instrument to maintain constant its metrological characteristics with time
- **Testing** is a technical investigation, e.g. as to whether a product fulfils its specified performance.
- **Traceability** means that a measured result can be related to stated references, usually national or international standards, through an unbroken chain of comparisons, all having stated uncertainties.
- **Trueness** - the closeness of agreement between the average value obtained from a large series of test results and an accepted reference value. The measure of trueness is usually expressed in terms of bias.
- **Uncertainty** of measurement is a parameter, associated with the result of a measurement, that characterizes the dispersion of the values that could reasonably be attributed to the measurand. It can also be expressed as an estimate characterizing the range of values within which the true value of a measurand lies.
- **Verification** is an investigation that shows that specified requirements are fulfilled.

3.0 TEST ACCURACY RATIO (TAR) MYTH

This is the most commonly abused requirement stated in calibration and test requirements by the industry. In simple terms, 10:1 TAR states that the standard used to calibrate the equipment has to be capable of achieving 10 times the accuracy.

It is easy to acquire a micrometer capable of measuring 0.00005 inches for less than US $ 100.00. In order to calibrate this micrometer, one would require a standard capable of measuring 0.000005 inches (implying 10:1 ratio). While this may be possible in this case, consider the other requirement for calibration, which is traceability. This means that the standard used has to have an unbroken chain of comparisons to a national standard (NIST in USA). This calibration traceability pyramid is illustrated in Figure 1.

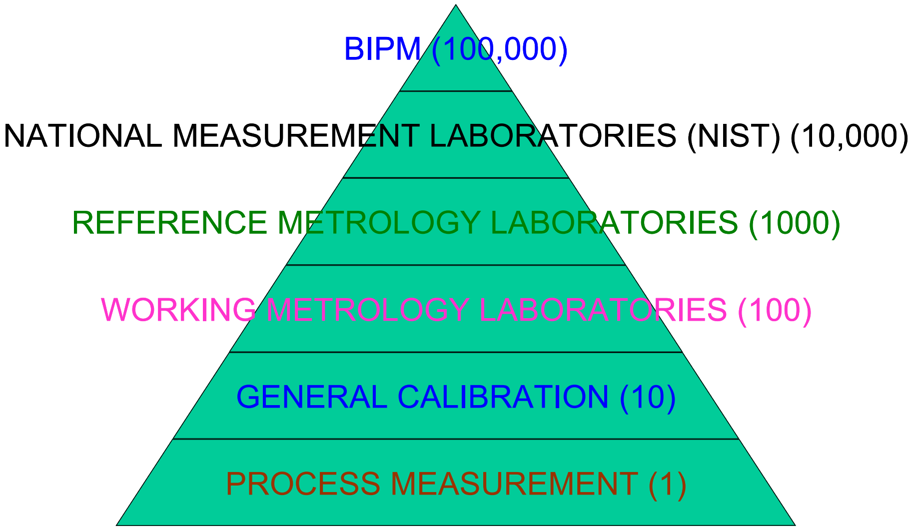

Figure 1: The 10:1 TAR Ratio and Calibration Traceability

As one moves up the calibration traceability pyramid, it becomes apparent that standards capable of achieving accuracy levels required are just not existent. The term accuracy is also misunderstood in defining this requirement. Sometimes the resolution of the instrument is the sole criteria defined in the accuracy of the instrument with no regard to other instrument characteristics that go into defining the accuracy specification.

To overcome this difficulty of achieving 10:1 TAR, the 4:1 TAR was introduced and judged acceptable. However, the 4:1 TAR becomes difficult to achieve as one moves up the traceability pyramid illustrated in Figure 2.

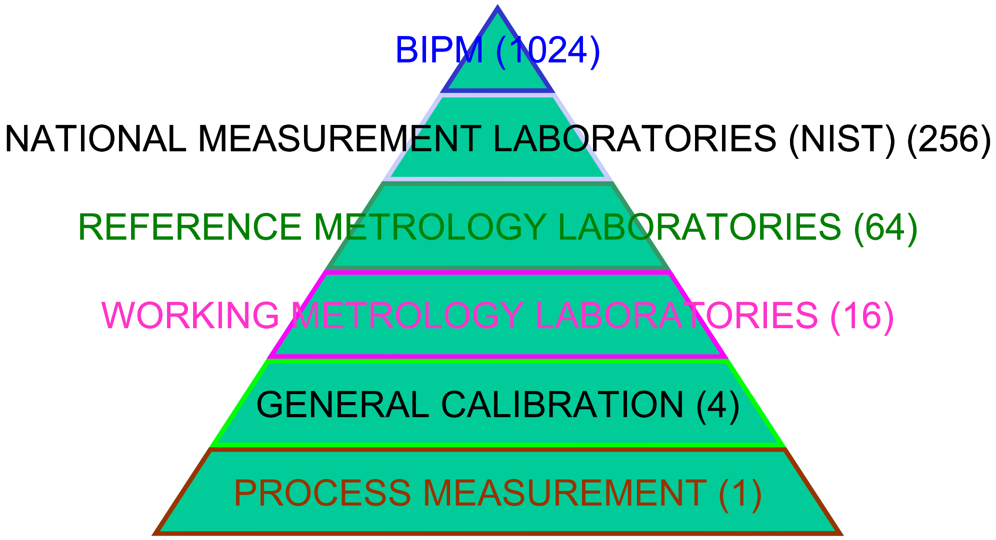

Figure 2: The 4:1 TAR Ratio and Calibration Traceability

Based on the definition of traceability, the uncertainties are cumulative. Thus the calibration traceability pyramid is now more realistic in making the measurement claims through use of estimated uncertainties. Used properly, the traceable calibration with measurement uncertainty is a more credible way to report the calibration data than making TAR claims.

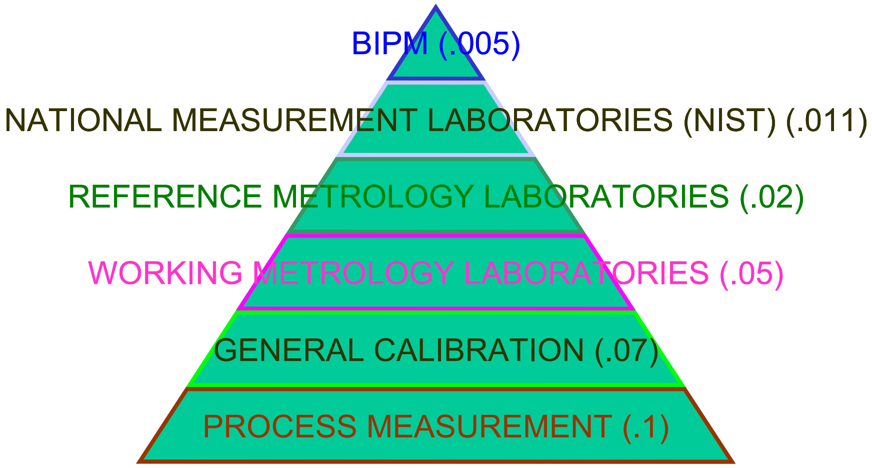

Figure 3: Calibration Traceability Pyramid

4.0 MEASUREMENT UNCERTAINTY

Other factors influence in making a measurement. These factors influence the overall measurement and complement in estimating the measurement uncertainty besides the use of manufacturer's stated specifications.

According to the GUM, the process of estimating measurement Uncertainty can be broken down in the following simplified steps. The author has identified several common errors observed while working with both the calibration laboratories and their customers.

1. **Identify the uncertainties in the measurement process.**

   The factors contributing to the uncertainty can come from a variety of sources. It is very common for the industry to ignore many factors affecting the measurement process. Environmental considerations, operator to operator or operator to equipment interactions are often ignored. Long term stability data is available but rarely included in the uncertainty budgets. Many laboratories do not put enough effort in this exercise.

2. **Classify type of uncertainty (A or B). Quantify (evaluate and calculate) individual uncertainty by various methods.**

   Type A evaluation method: The method of evaluation of uncertainty of measurement by the statistical analysis of series of observations.

   Repeatability and Reproducibility analysis can be easily determined with computer software tools yet this information is not included in the uncertainty budget considerations. Analysis techniques such as Statistical Process Control (SPC), Analysis of Variance (ANOVA) and Design of Experiments (DOE) are not utilized as often.

   Type B evaluation method: The method of evaluation of uncertainty of measurement by means other than the statistical analysis of series of observations.

   While more estimates of Type B uncertainty components are included in the budget, Type B estimates of uncertainty are not sufficiently analyzed. When included in the uncertainty budgets, standard uncertainty is not derived using the familiar correction factors depending on rectangular, triangular or U-shaped distributions.

3. **Document in an uncertainty budget.**

   Uncertainty budgets do not contain enough information to determine how the estimates are derived. Once the budgets are developed, they get filed and are not reviewed as often when the calibration process changes or new equipment is added. It is difficult to second guess how the information was derived if sufficient information (including raw data) is not included in budget.

4. **Combine uncertainty (Root Sum Square (RSS) method).**

$$u_{c_a} = \sqrt{u_{c_{a1}}^{\;2} + u_{c_{a2}}^{\;2} + \ldots\ldots}$$ Combine Type A uncertainties

$$u_{c_b} = \sqrt{u_{c_{b1}}^{\;2} + u_{c_{b2}}^{\;2} + \ldots\ldots}$$ Combine Type B uncertainties

$$u_c = \sqrt{u_{c_a}^{\;2} + u_{c_b}^{\;2}}$$ Combine all Type A and Type B uncertainties

The root sum square method is used for combining all the uncertainty components. No determination is made for correlated components of uncertainty. In most cases, this does not penalize the uncertainty budget as much (remember: it is an estimate), in some critical cases, it does misrepresent data to make a difference.

5. **Assign appropriate k factor multiplier to combined uncertainty to report expanded uncertainty.**

$U = k \cdot u_c$ The GUM specifies k = 2 or 95% confidence interval.

Although, this is a straight multiplier, industry in general does not understand the use of confidence intervals and degrees of freedom. Sometimes, misstating uncertainty at k=1 makes the numbers look good on the scope of accreditation. Various numbers games are played using uncertainty numbers in what is called the "Scope Wars" among commercial calibration laboratories. It is the responsibility of the accrediting body to ensure that a laboratory's scope of accreditation is stated accurately. However, there is misrepresentation of data in the scope of accreditation.

5.0 OTHER INDUSTRY MISAPLICATION OF MEASUREMENT UNCERTAINTY ESTIMATES

From the calibration service user's point of view, it is easier to state that the measurement equipment is calibrated with its "badge of honor" (namely the calibration sticker) than understand the measurement uncertainty and its proper application. The uncertainty stemming from the industry in not knowing how measurement uncertainty estimates are applied in their daily application results in various misapplications of measurement uncertainty. They are listed below.

1. Equipment with larger uncertainty than the required tolerance is used to make measurements.
2. Equipment with very small uncertainty is used where required tolerance is large.
3. Judgment is based on a single measurement on a critical parameter.
4. Lack of compensation for environment when measurements are made.
5. More significant digits on measurement equipment mean more accuracy.
6. Making a measurement with equipment with less resolution and increasing the precision (i.e. more decimal places) by manipulating the numbers in a spreadsheet.
7. Claiming very low uncertainties on measuring equipment incapable of resolving such low values.
8. Lack of a measurement assurance program to ensure that the estimated measurement uncertainty has not changed and assuring confidence in the measurement process over a period of time.

6.0 SUMMARY

This paper outlined some misapplications and errors sources contributing to the uncertainty of the (pun intended) application of measurement uncertainty by the both the industry and calibration laboratories alike. The GUM is a great reference document for the experienced metrologist who is also a statistician. However, it is not an easy document to read for a common practitioner or a user of calibration services. Guide supplements to GUM for the user of calibration services (Part A) and practitioner (Part B) should help resolve this confusion. As we move toward a very global economy, the one common force that drives the world commerce will the measurement process. Quantifying and understanding the uncertainties associated with the measurement become even more important as we apply it to other non-traditional uses besides metrology.

7.0 REFERENCES

1. Group, Z540-2, 1997, *U. S. GUIDE TO EXPRESSION OF UNCERTAINTY IN MEASUREMENT,* Boulder, ANSI/NCSL International
2. ISO 1993. International vocabulary of basic and general terms in Metrology (VIM), Geneva: ISO
3. ANSI/ISO 17025-1999, American National Standard-General requirements for the competence of testing and calibration laboratories. Milwaukee: ASQ Quality Press
4. Bucher, Jay L. 2004. The Metrology Handbook. Measurement Quality Division. Milwaukee: ASQ Quality Press