Measurement Uncertainty and Validation

Table of Contents

Measurement Uncertainty and Validation

1 PURPOSE

This document describes the method for evaluating and estimating the measurement uncertainty and the selection, verification, and validation of methods that are specific to the Firearms/Toolmarks Discipline (FTD). The FTD is composed of personnel from the Firearms/Toolmarks Unit (FTU) and the Scientific & Biometrics Analysis Unit – Toolmark Group (SBAU-TG).

2 SCOPE

This procedure applies to FTD personnel utilizing technical procedures wherein a quantitative measurement is reported, such as barrel and overall length measurements. Additionally, this procedure applies to FTD personnel when the measurement uncertainty is requested by the contributor. This procedure does not apply to approximate ranges, such as those reported in gunshot residue examinations and shooting incident reconstructions.

3 MEASUREMENT UNCERTAINTY EVALUATION

The result of a measurement is an approximation or estimate of the value of the specific quantity subject to measurement (measurand) and is only complete when accompanied by a quantitative statement of its uncertainty. The following steps describe the method for estimating the measurement uncertainty within the FTD and are based on the National Institute of Standards and Technology's (NIST) eight-step process for estimating uncertainty.

3.1 Estimating Measurement Uncertainty

A. Estimating the uncertainty for quantitative procedures follows the NIST eight-step process:
    1. Define what is being measured
    2. Identify sources of uncertainty
    3. Quantify uncertainty sources
    4. Convert factors to standard uncertainties
    5. Calculate combined standard uncertainties
    6. Expand the uncertainty by coverage factor (k)
    7. Evaluate the expanded uncertainty
    8. Report results with uncertainty

B. Uncertainty estimates involve the development of an uncertainty budget. The uncertainty budget for a procedure will include both Type A uncertainties which are directly calculated from repeated measurements of some quantity and Type B uncertainties which are estimated using other information such as instrument calibration reports provided by a vendor.

C. An uncertainty budget will be recorded on the FTD Uncertainty Budget Form and the appropriate Unit Quality Assurance Representative will ensure that a copy is maintained and made available.

3.1.1 *Define what is being measured (measurand)*

The Examiner will state what is being measured on an FTD Uncertainty Budget Form, which includes information about the procedure, date, sources of uncertainty, and name of preparer(s).

3.1.2 *Identify sources of uncertainty (budget)*

A. The Examiner must first attempt to identify sources of uncertainty associated with the process of measuring. It is recognized that identifying "all" uncertainty components which contribute to the measurement uncertainty may not be achievable. In order to identify possible sources of uncertainty when collecting a measurement, the following potential sources of uncertainty should be considered. Not all will be applicable in every measurement situation.

1. Sampling
    - i. Homogeneity
    - ii. Effects of specific sampling strategy (e.g., random, stratified random, proportional)
    - iii. Temperature and pressure
2. Sample preparation
    - i. Homogenization and/or sub-sampling effects
    - ii. Contamination
3. Presentation of Certified Reference Materials (CRM) to measuring system
    - i. Uncertainty of CRM
    - ii. CRM match to sample
4. Calibration of Instrument
    - i. Instrument calibration errors using CRM
    - ii. Reference material and its uncertainty
    - iii. Instrument precision
5. Analysis (data acquisition)
    - i. Operator effects
    - ii. Instrument parameter settings
    - iii. Run-to-run precision
6. Data processing
    - i. Averaging
    - ii. Effects of rounding and truncating
    - iii. Statistics
    - iv. Processing algorithms (model fitting, e.g., linear least squares)
7. Presentation of results
    - i. Final results
    - ii. Estimation of uncertainty
    - iii. Confidence level
8. Interpretation of results
    - i. Against limits/boundaries
    - ii. Regulatory compliance

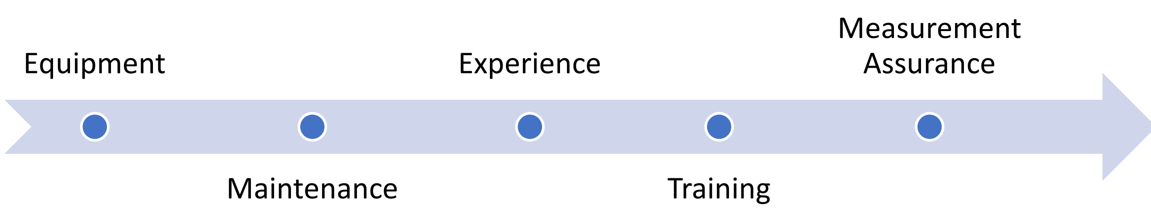
iii. Fitness for purpose

B. Another way to identify and sort sources of uncertainty is to use a cause and effect diagram. This method helps to visualize how the different sources of uncertainty relate to one another and to reduce the possibility of counting the same source of uncertainty more than once.

3.1.2.1 Cause and Effect Diagram

The initial step in creating a cause and effect diagram involves brainstorming the main sources of uncertainty. This determines the main branches of the diagram. Next, consider each step of the method and add any additional factors to the diagram. For each branch, add contributing factors. Do not add contributing factors that you believe to be remote possibilities. Resolve the duplications as outlined below:

- *Cancelling Effects* - when the source's effect is shown to have no net effect on the result, it can be cancelled/removed from the diagram.
- *Same effect/same time* - when a similar source is uncovered on multiple branches of a cause and effect diagram, they can be combined into a single source happening at the same time. A common occurrence is that reproducibility appears on many branches; these can be combined into a single source.
- *Different Instances/Re-labeling* - if there are similarly named effects that refer to different instances of similar measurements, these effects will be re-labeled to clearly distinguish them from each other.



3.1.2.2 Reconciliation of Uncertainty Components

The process of reconciliation simplifies the uncertainty budget. In this step, a review is conducted to determine whether a listed uncertainty source is adequately accounted for by the existing data. The basis for this step lies in the fundamental assumption that if an effect is representatively varied during the course of a series of observations, the uncertainty associated with that effect is adequately accounted for in the standard deviation of those observations.

3.1.3 <u>Quantify uncertainty sources</u>

Once the sources of uncertainty have been established, the Examiner must measure or estimate the magnitude of the uncertainty. Once the available information has been collected, each potential source of uncertainty is evaluated and categorized as Type A or Type B uncertainty data.

- Type A is a method of evaluation of uncertainty by statistical analysis of a series of observations.
- Type B is a method of evaluation of uncertainty by means other than the statistical analysis of a series of observations.

3.1.3.1 Type A uncertainties

The inability to exactly reproduce all parameters of a measurement, combined with precision limits of measurement devices, leads to measurement values being randomly dispersed. This random dispersion of measured values is referred to as Type A uncertainty. Type A uncertainties are estimated using statistics from repeated measurements.

$$S\ (\text{Sample}) = \sqrt{\left(\sum_{i=1}^{n}(Xi - \overline{x})^2/(n-1)\right)}$$

The random uncertainty of a population is determined by evaluating the standard deviation of the mean. This is obtained by dividing the standard deviation of the sample by √ (number or "n").

$$\sigma\ (\text{Standard deviation of the mean}) = S/\sqrt{n}$$

This is the Type A uncertainty: $U_{A1} = S/\sqrt{n}$. If there is more than one contributor to the Type A uncertainties, repeat the above process for each one of them. Then the random uncertainty is $U_A = \sqrt{(U^2_{A1} + U^2_{A2} + U^2_{A3}.....)}$. Prior to combining factors, it is important to ensure that all values are expressed in the same units.

3.1.3.2 Type B uncertainties

Type B uncertainties occur due to sources of uncertainty and bias in a measuring system and/or method that are evaluated by means other than direct statistical evaluation. These uncertainties can be minimized by optimizing the design of a measuring system and/or method to reduce their contribution. Additionally, it is acceptable for systematic uncertainties to be estimated. Systematic uncertainties are those where the same influence affects the result for each repeated measurement. Although these factors may contribute insignificantly to the overall uncertainties, they need to be considered for Type B evaluation. When evaluating Type B sources of uncertainty for a measurement process, the list below highlights the more common sources:

- *Equipment* - The equipment chosen to conduct measurements will be NIST traceable and calibrated at a minimum to the manufacturer(s)' designated calibration intervals to reduce the source of uncertainty.
- *Personnel* - The differing physical capabilities, experience and abilities of the personnel performing a given measurement can affect the observed values. These contributions to the uncertainty may be treated as a Type B effect.
- *Readability* - Readability is defined as the smallest increment that can be detected by the measuring equipment. Due to its typically small contribution to the uncertainty, readability is usually not considered to be an issue relative to the practical certainty involved with the measurements.

- *Calibration/Calibration Bias* - The uncertainty associated with calibration is located on the calibration certificate. The uncertainty associated with calibration is expressed as an expanded uncertainty and is assumed to be a normal distribution.
- *Facility/Environment* - The environmental conditions in the Laboratory are typically maintained at a relatively constant conditions. However, some evaluation is required to verify their effects have been appropriately accounted for.
- *Pressure* - When taking dimensional measurements the effect of applied pressure from the instrument on the measurand should be considered.

Many Type B uncertainty contributions can be determined using sources of information such as calibration certificates, reference data, and manufacturer(s) specifications. Such contributions include:

- Uncertainty contributed due to the deviation of the reference standard from its nominal value. (Assume a rectangular distribution and a coverage factor of $k=\sqrt{3}$)
- Uncertainty due to the calibration of reference standard. (The coverage factor is obtained from the certificate and is usually $k=2$, for normal distribution)
- Uncertainty contribution due to resolution of the unit under calibration (½ resolution and $k=\sqrt{3}$)
- Uncertainty contribution due to the resolution of the temperature measuring device (½ resolution and assume distribution is rectangular so that $k=\sqrt{3}$)
- Uncertainty contribution due to uncertainty of the temperature measuring device (this value and the coverage factor, k, are obtained from the calibration certificate)

When these systematic uncertainties are unavailable, they can be estimated using the experience or general knowledge of the behavior and properties of relevant materials and instruments.

Calibration certificates which are generated by NIST or another accredited laboratory with traceability to a NIST standard will typically provide a 95% confidence level (2σ). This must be compensated for in the systematic uncertainty calculations (divide by 2).

If a non-accredited laboratory or manufacturer's specifications are used to determine the systematic uncertainty of a measurement or an estimate must be made outside the limits of the uncertainty of a measurement, a rectangular distribution should be assumed. With a rectangular distribution, the range of the outer limit (2a) is used to estimate the standard deviation using the $\sigma = a/\sqrt{3}$.

3.1.4 Convert factors to standard uncertainties

Standard deviation is also known as the estimated standard uncertainty. In this step, all previous standard deviations are expressed as standard uncertainties. However, to facilitate this step, it is necessary that common units are used throughout the budget. If this is not possible, conversion of units into percentages (i.e., relative standard uncertainty) is necessary; keeping in mind that the reverse conversion will be necessary later in the process.

3.1.5 *Calculate combined standard uncertainties*

The individual standard uncertainties quantified by Type A and Type B evaluations are now combined to calculate the combined standard uncertainty. The Root Sum Square technique is used to calculate the combined standard uncertainty ($U_{combined}$ or U_c) which is expressed as follows: $U_c = \sqrt{(U_{SD}^2 + U_{resolution}^2 + U_{calibration}^2 + U_{tempcoef}^2 \ldots)}$.

3.1.6 *Expand the uncertainty by coverage factor (k)*

The coverage factor (k) is a number that, when multiplied by the combined standard uncertainty ($U_{combined}$), produces an interval around the average measurement result that is expected to include a large specified percentage (usually 95% or 99.7%) of the values. Usually k is set to a value of 2 to represent 95% confidence level and 3 to represent 99.7% confidence level. Within the FTD, a confidence level of 99.7% or greater is used in reporting a quantitative result.

When the Type A uncertainty component is dominant and the number of measurements used to calculate the standard deviation is less than 100, there is a reduced confidence in the calculated standard deviation. Since a normal distribution model indicates that the results close to the mean are more probable than results far from the mean, when only a few measurements are made, it is likely an underestimation of the true standard deviation. To account for this, a correction factor can be applied based on the Student t-distribution. The following table shows the corrected values for k as a function of degrees of freedom (n-1) where n is the number of measurements (*For the FTD the minimum number of measurements is ten*).

n-1	$k_{correction}$	n-1	$k_{correction}$	n-1	$k_{correction}$	n-1	$k_{correction}$
1	212.2	9	4.02	17	3.46	50	3.12
2	18.22	10	3.89	18	3.43	60	3.09
3	8.89	11	3.79	19	3.40	70	3.08
4	6.44	12	3.71	20	3.38	80	3.06
5	5.38	13	3.64	25	3.29	90	3.05
6	4.80	14	3.58	30	3.23	100	3.04
7	4.44	15	3.54	35	3.19	∞	3.0
8	4.20	16	3.49	40	3.16		

Student-t Table distribution for k correction factor at a 99.7% Confidence Level

Finally, if standard uncertainties were converted to percentages, the final combined uncertainty should be converted back to the units of measurement of the material in question.

3.1.7 *Evaluate the expanded uncertainty*

The estimation uncertainty determined from the preceding steps will be evaluated to determine if it "makes sense" and is "reasonable" for the procedure being evaluated.

3.1.8 *Reporting results with uncertainty*

A. The coverage probability will be no less than the approximated 99.7% confidence level and the numerical value of the expanded uncertainty will be reported to at most two significant digits, unless there is a recorded rationale for reporting additional significant digits.

B. When reporting the uncertainty of a measurement, rounding the uncertainty upwards rather to the nearest digit will be performed.
C. The associated expanded uncertainty will be reported to the same level of significance as the measurement result.
D. Measurement uncertainty and confidence level will be included in a Laboratory Report, or as an enclosure, when it is relevant to the validity or application of the examination results or a customer's instructions require it. For example, a report may state the barrel length of the Item 13 shotgun was measured to be 16.56" ± 0.07" (k =3 for 99.7% confidence level).

4 SELECTION, VERIFICATION, AND VALIDATION OF METHODS

4.1 Validation Study

Minimum requirements for a validation study in the FTD will include:

- Identifying the limitations of the procedure, reported results, opinions, and interpretations.
- Conditions under which reliable results can be obtained.
- Critical aspects of the procedure that must be controlled and monitored.
- The scope and accuracy of the procedure to meet the needs of the given application.
- The associated data analysis and interpretation.
- Establishing the data required to report a result, opinion, or interpretation.

4.2 Validation Reviews

A. Approvals obtained for records of method selection, development, validation, to include software validations, and competency tests may be recorded electronically.

B. Documentation of these approvals will be maintained within the approved plan and include the version of the plan reviewed, name of approver, and the date of approval.

5 REFERENCES

Barry N. Taylor and Chris E. Kuyatt (2001). *Guidelines for Evaluating and Expressing the Uncertainty of NIST Measurement Results*, [Online]. Available: http://physics.nist.gov/TN1297 [February 2022]. National Institute of Standards and Technology, Gaithersburg, MD.

Stephanie Bell (2001). *Measurement Good Practice Guide No. 11* (Issue 2), [Online]. Available: http://www.demarcheiso17025.com/private/A%20Beginner%92s%20Guide%20to%20Uncertainty%20of%20Measurement%20.pdf [February 2022]. National Physical Library, Teddington, Middlesex, United Kingdom.

Guide for Evaluation of Uncertainty in Calibration (2008), [Online]. Available: https://www.iasonline.org/resources/guide-for-evaluation-of-uncertainty-in-calibration/ [February 2022]. International Accreditation Service, Inc., Brea, CA.

Quantifying Uncertainty in Analytical Measurement, 3rd Edition (2012), [Online]. Available: https://www.eurachem.org/index.php/publications/guides/quam [February 2022]. EURACHEM/CITAC Measurement Uncertainty Working Group.

Joint Committee for Guides in Metrology (JCGM), *Evaluation of measurement data – Guide to the expression of uncertainty in measurement* (GUM) (GUM 1995 with minor corrections). (Serves, France: International Bureau of Weights and Measures [BIPM]-JCGM 100], September 2008).

https://services.anu.edu.au/research-support/tools-resources/statistical-software-packages (last visited 02/12/2022)

6 Revision History

Revision	Issue Date	Changes
00	2/18/2022	Original issuance of document. Transferred information from FTD-106-04 and FTD-108-02 and updated to new LOM and QAM requirements.