Forensic Metrology

A Primer for Lawyers and Judges

Ted Vosk

Forensic Metrology
A Primer for Lawyers and Judges

2nd Edition

Supplemental materials for the textbook:
Vosk, T., and Emery, A., *Forensic Metrology: Scientific Measurement and Inference for Lawyers, Judges, and Criminalists* (CRC/Taylor Francis Group – Available Sept. 2014)

Ted Vosk

FIRST PUBLISHED FOR
Forensic Metrology: The Key to the Kingdom
National Forensic Blood and Urine Testing Seminar
Georgia Association of Criminal Defense Lawyers
San Diego CA, May 14, 2009

V 1.10

8105 NE 140th Pl., Bothell WA 98011

FOR KRIS

8105 NE 140th Pl., Bothell WA 98011

Table of Contents

8105 NE 140th Pl., Bothell WA 98011

WHY FORENSIC METROLOGY FOR LAWYERS AND JUDGES?[1]

Over the past decade the forensic sciences have come under increasing fire by scientists and legal professionals alike, culminating in the recent National Academy of Sciences report: *Strengthening Forensic Science in the United States: A Path Forward*. Many, if not most, forensic scientists are dedicated professionals. Nonetheless, burgeoning caseloads, pressure to assist prosecutions and a lack of resources have led to systemic failures to adhere to basic scientific standards, principles and practices. Given the significant role scientific knowledge and evidence plays in the courtroom, these weaknesses threaten to undermine the integrity of our system of justice as a whole. It is beyond question that the current state of forensic science needs widespread reform and forensic scientists have bore a brunt of the blame.

Forensic scientists are only one side of the coin, however. Sharing equal blame for this state of affairs are lawyers and judges who encounter forensic science in the courtroom on an increasingly frequent basis. Many of these professionals expend great effort to understand and critically assess forensic practices. Unfortunately, many do not. Frequent is the refrain from lawyer and judge alike that the reason they went to law school was so that they wouldn't have to do science or math anymore. Of those lawyers and judges who do endeavor to gain an understanding of forensic matters before them, many become overwhelmed by complexities and the matter of even knowing where to begin. Uncritical acceptance, "science-phobia" and even lethargy have lead to frequent reliance upon evidence that isn't even good enough to be called wrong.[2] Thus, if the integrity of our justice system is to be preserved, it is equally important to reform the legal practices of lawyers and judges. In today's

[1] This primer is based on the textbook: Vosk, Emery, Fitzgerald, *Forensic Metrology: A Primer on Scientific Measurement for Lawyers, Judges and Forensic Scientists* (CRC Press – In Preparation).

[2] The phrase indicating that scientific work that is so poorly done that it isn't even good enough to be accorded the status of being called wrong is attributable to Nobel Physicist Wolfgang Pauli. Gieser, *The Innermost Kernel; Depth Psychology and Quantum Physics. Wolfgang Pauli's Dialogue with C.G. Jung* 72 (Springer 2005).

8105 NE 140th Pl., Bothell WA 98011

technologically advanced society, the law's truth finding function cannot be achieved if its practitioners are ignorant of the basic tenants of science.

The foundation of all science is measurement and observation. Measurement and observation act as both the genesis of scientific understanding and/or confirmation for theory based models. Absent these activities, the only thing that binds our notions of the physical world to reality is faith.[3] We would be left to speculate about, and ponder upon, the workings of nature but with little reason to place confidence in our conclusions. Systematic measurement and observation, objectively and without malice, demonstrate where our physical notions are wrong. But they also reveal to us regularities fundamental to the physical world, permitting us to build models for purposes of predicting how it will behave in a given set of circumstances. Rigorous, systematic measurement and observation are necessary to the acquisition and proper application of all scientific knowledge.

This understanding is equally critical for those who rely upon, or engage in the application of, science in matters of common daily interest. Whether it's weighing out the proper proportions of medication for a prescription at the corner pharmacy, interpreting the results of a pregnancy test or determining how fast an individual's automobile is traveling through a speed zone, each relies upon measurement and observation. Certainly the degree of rigorousness required depends on the importance we place on the correctness of the determinations being made. The point to be illustrated, however, is simply that proper measurement and observation lie at the foundation of all scientific determinations, even when not recognized as such, regardless of the field of investigation or application.

This leads to an astonishing conclusion. If there are principles of measurement and observation that remain fundamental regardless of application, they provide a discrete tool for critical evaluation of certain aspects of *all* scientific claims, even absent expertise in a specific area.

[3] The author is not disparaging faith as an equally valid way of knowing and understanding the world. It is simply a matter that science and faith are distinct approaches, the former involving belief based on proof, the latter involving belief even in the absent of proof. Since in the courtroom it is proof rather than faith that must determine belief, scientific evidence is what is relied upon.

8105 NE 140th Pl., Bothell WA 98011

Metrology,[4] the science of measurement and observation, provides such principles and tools. Metrological[5] principles apply to every measurement and observation made in every lab anywhere on the planet. As noted physicist Lord Kelvin said over a century ago, "…if science is measurement, then without metrology there can be no science."[6] Thus, given a basic understanding of metrology, even a nonscientist can begin to engage in a critical analysis of scientific claims across a broad spectrum based on metrological principles. For the lawyer or judge, this means that even in the absence of an expert, he/she can engage in a critical analysis of whether forensic science being presented is metrologically sound.

One might wonder about the scope of these metrological tools. The report issued by the National Academy of Sciences focuses to a large extent on the metrological failures of the forensic sciences. These include: a lack of methodological standards; the failure to determine, understand or report the inherent limitations and uncertainty associated with methods and results; and the absence of mechanisms ensuring quality control of methods utilized. Much of what was reported in the National Academy report could have been discovered by anybody with a modest understanding of metrology and its application to the forensic sciences.

So we return to the question from which we started: Why forensic metrology for lawyers and judges? By gaining a basic understanding of metrological principles, judges and lawyers gain a basic understanding of science itself. In the courtroom, they are empowered to become participants in the critical analysis of forensic evidence instead of passive, intimidated or confused spectators. Outside the courtroom, they become voices of informed reason to help shape scientifically sound forensic policy. Most importantly, though, armed with a better understanding of the scientific process, they help preserve the integrity of our system of justice and facilitate it's ultimate goal of determining truth in the matters subject to it.

[4] Including the emerging field of "proto-metrology".
[5] Including the emerging field of "proto-metrology".
[6] William Thompson (Lord Kelvin), *Lecture to the Institution of Civil Engineers*, May 3, 1883.

8105 NE 140th Pl., Bothell WA 98011

I. CONCERNING FORENSIC SCIENCE AND THE ADMINISTRATION OF JUSTICE

A. SCIENCE.

1. "Appropriate scientific standards are widely ignored in forensic laboratories"[7] contributing "to questions about the validity of conclusions."[8]

2. "Few forensic science methods have developed adequate measures of the accuracy of inferences made by forensic scientists."[9]

3. "Much forensic evidence…is introduced in criminal trials without any meaningful scientific validation, determination of error rates, or reliability testing."[10]

4. "[F]orensic scientists themselves often fail to consider or appreciate measurement uncertainty."[11]

5. "The process leading from evidence to conclusion is often opaque, either because it lacks scientific rigor and is inherently unfalsifiable, or because the approach is inadequately tested, and thus cannot quote random match probabilities or estimate the chance of error."[12]

6. "There is a critical need in most fields of forensic science to raise the standards for reporting and testifying about the results of investigations…imprecision in vocabulary stems in part from the paucity of research in forensic science and the corresponding limitations in interpreting the results of forensic analyses."[13]

7. "[B]ad laboratory practices have bedeviled even the FBI laboratories."[14]

8. "Criminal justice agencies have been slow to adopt new scientific procedures…despite repeated calls for accreditation and oversight, many government crime labs continue to lack either one…justice would be furthered by a more scientific and reliable technology for analyzing crimes. The mystery here is why the practitioners don't seem to want it!"[15]

9. "THE LAW'S GREATEST DILEMMA IN ITS HEAVY RELIANCE ON FORENSIC EVIDENCE, HOWEVER, CONCERNS THE QUESTION OF WHETHER—AND TO WHAT EXTENT—THERE IS *SCIENCE* IN ANY GIVEN 'FORENSIC SCIENCE' DISCIPLINE."[16]

---

[7] Erica Beecher-Monas, *Evaluating Scientific Evidence: An Interdisciplinary Framework for Intellectual Due Process*, 97 (Cambridge Press 2007).
[8] NAS, *Strengthening Forensic Science in the United States: A Path Forward*, 7-7 (2009).
[9] NAS, *Strengthening Forensic Science in the United States: A Path Forward*, 6-1 (2009).
[10] NAS, *Strengthening Forensic Science in the United States: A Path Forward*, 3-18 (2009).
[11] Gullberg, *Estimating the measurement uncertainty in forensic breath-alcohol analysis*, 11 ACCRED. QUAL. ASSUR. 562, 563 (2006).
[12] Gonzalez-Rodriguez, *Emulating DNA: Rigorous Quantification of Evidential Weight in Transparent and Testable Forensic Speaker Recognition* 15(7) IEEE Transactions on Audio, Speech, and Language Processing 2104, 2104 (2007).
[13] NAS, *Strengthening Forensic Science in the United States: A Path Forward*, 6-3 (2009).
[14] Erica Beecher-Monas, *Evaluating Scientific Evidence: An Interdisciplinary Framework for Intellectual Due Process*, 97 (Cambridge Press 2007).
[15] Kennedy, *Forensic Science: Oxymoron*?, 302 Science 1625, 1625 (2003).
[16] NAS, *Strengthening Forensic Science in the United States: A Path Forward*, 3-2 (2009).

8105 NE 140th Pl., Bothell WA 98011

B. LAWYERS AND JUDGES.

1. "The judicial system is encumbered by…judges and lawyers who generally lack the scientific expertise necessary to comprehend and evaluate forensic evidence in an informed manner."[17]

2. "Defense attorneys, prosecutors, judges and lay juries often lack scientific training and naively accept measurement results as certain."[18]

3. "[L]awyers…do not know how to think about validation of science claims."[19]

4. "It is difficult to persuade a judge or a court that there is no certainty in measurement results…Yet, considering or not the uncertainty of a critical result can make the difference between acquittal and a guilty sentence."[20]

5. "[E]stablished case law in many jurisdictions supports minimal analytical quality control and documentation."[21]

6. "[L]egislators, government officials, judges, lawyers, and juries are not noted for their technical literacy, let alone their understanding of the intricacies of metrology in chemical and measurement uncertainty."[22]

7. "MANY LAWYERS SIMPLY COULD NOT DISTINGUISH BETWEEN REAL SCIENCE AND PRETENSIONS TO SCIENCE."[23]

II. WHERE TO BEGIN?

A. SCIENCE 101

1. "Scientific method refers to the body of techniques for investigating phenomena, acquiring new knowledge, or correcting and integrating previous knowledge. It is based on gathering observable, empirical and measurable evidence subject to specific principles of reasoning."[24]

   a. "Measurement…is the essential tool by which humans describe the world and reason about it."[25]

---

[17] NAS, *Strengthening Forensic Science in the United States: A Path Forward*, 3-20 (2009).
[18] Gullberg, *Estimating the measurement uncertainty in forensic breath-alcohol analysis*, 11 ACCRED. QUAL. ASSUR. 562, 563 (2006).
[19] Saks, *Failed Forensics: How Forensic Science Lost Its Way and How It Might Yet Find It*, 4 ANNU. REV. LAW SOC. SCI. 149, 153 (2008).
[20] Bich, *Interdependence between measurement uncertainty and metrological traceability* ACCRED. QUAL. ASSUR. (IN PRESS - 2009).
[21] Gullberg, *Estimating the measurement uncertainty in forensic breath-alcohol analysis*, 11 ACCRED. QUAL. ASSUR. 562, 563 (2006).
[22] King, *Chemical measurement and the law: metrology and quality issues*, 6 ACCRED. QUAL. ASSUR. 236, 243 (2001).
[23] Saks, *Failed Forensics: How Forensic Science Lost Its Way and How It Might Yet Find It*, 4 ANNU. REV. LAW SOC. SCI. 149, 153 (2008).
[24] Sir Isaac Newton, *Philosophiae Naturalis Principia Mathematica* (1687).
[25] Finkelstein, *Expanding Technology, Deepening Knowledge and a Shrinking World: Reflections on Learned Societies in Measurement and Instrumentation*, 41 MEAS. CONTROL 170, 170 (2008).

8105 NE 140th Pl., Bothell WA 98011

b. If "measurements are flawed, analyses and interpretations based on these measurements are fundamentally and irreparably fallacious."[26]

2. "MEASUREMENT THEORY IS THE CONCEPTUAL FOUNDATION OF ALL SCIENTIFIC DECISIONS."[27]

B. METROLOGY: "Science of measurement and its application. Metrology includes all theoretical and practical aspects of measurement, whatever the measurement uncertainty and field of application."[28]

1. "Metrology is multi-disciplinary…In application, metrology enables measurements of potentially all quantities to be related to one another in a true and absolute sense – that is the key of metrology."[29]

2. "…IF SCIENCE IS MEASUREMENT, THEN WITHOUT METROLOGY THERE CAN BE NO SCIENCE."[30]

III. METROLOGY: A LITTLE BACKGROUND

A. METROLOGIST: "Develops and evaluates calibration systems that measure characteristics of objects, substances, or phenomena, such as length, mass, time, temperature, electric current, luminous intensity, and derived units of physical or chemical measure: Identifies magnitude of error sources contributing to uncertainty of results to determine reliability of measurement process in quantitative terms. Redesigns or adjusts measurement capability to minimize errors. Develops calibration methods and techniques based on principles of measurement science, technical analysis of measurement problems, and accuracy and precision requirements. Directs engineering, quality, and laboratory personnel in design, manufacture, evaluation, and calibration of measurement standards, instruments, and test systems to ensure selection of approved instrumentation. Advises others on methods of resolving measurement problems and exchanges information with other metrology personnel through participation in government and industrial standardization committees and professional societies."[31]

B. BRIEF HISTORY:

1. Bible: "Have true scales, true weights and measures for all things."[32]

2. Magna Carta: "There shall be standard measures of wine, ale, and corn (the London quarter), throughout the kingdom. There shall also be a standard width of dyed cloth, russett, and haberject, namely two ells within the selvedges. Weights are to be standardized similarly."[33]

3. U.S. Matters of State: "WEIGHTS AND MEASURES may be ranked among the necessaries of life to every individual of human society. They enter into the economical arrangements and daily concerns of every family. They are necessary to every occupation of human industry; to the

[26] Krebs, *Measurement Theory*, 67(12) PHYS. THERAPY 1834, 1839 (1987).
[27] Krebs, *Measurement Theory*, 67(12) PHYS. THERAPY 1834, 1834 (1987).
[28] JCGM, *International Vocabulary of Metrology — Basic and General Concepts and Associated Terms (VIM)*, § 2.2 (2008).
[29] Pendrill, *Metrology: time for a new look at the physics of traceable measurement?* 37(1) Europhysics News 24 (2006); Regtien, *Metrology as part and parcel of training programs for science and engineering*, 7(1) MEAS. SCI. REV. 9, 9 (2007).
[30] William Thompson (Lord Kelvin), *Lecture to the Institution of Civil Engineers*, May 3, 1883.
[31] U.S. Dept. of Labor, *Dictionary of Occupational Titles* 012.067-010.
[32] Leviticus 19:36.
[33] Magna Carta § 32.

8105 NE 140th Pl., Bothell WA 98011

distribution and security of every species of property; to every transaction of trade and commerce; to the labors of the husbandman; to the ingenuity of the artificer; to the studies of the philosopher; to the researches of the antiquarian; to the navigation of the mariner and the marches of the soldier; to all the exchanges of peace, and all the operations of war. The knowledge of them, as in established use, is among the first elements of education, and is often learned by those who learn nothing else, not even to read and write. This knowledge is riveted in the memory by the habitual application of it to the employments of men throughout life."[34]

IV. METROLOGY

A. METROLOGICAL FOCUS

1. MEASUREMENT: Process of experimentally obtaining one or more ***quantity values*** that can reasonably be attributed to a ***quantity***. Measurement does not apply to nominal properties.[35]

2. OBSERVATION (EXAMINATION): The process of obtaining information regarding the presence or absence of an attribute of a test specimen, or of making a reading on a characteristic or dimension of a test specimen.[36] Observation (examination) produces ***qualitative results*** indicating nominal and ordinal properties such as classification, identification and ordering.[37]

3. Traditionally metrology has been limited to measurements yielding quantitative results. In recent years the field of proto-metrology has developed to address observations yielding qualitative results.[38] I refer to them both under the common heading metrology herein for ease except where making a clear distinction is necessary.

B. WEIGHTS & MEASURES

1. MEASUREMENT UNIT: Real scalar quantity, defined and adopted by convention, with which any other quantity of the same kind can be compared to express the ratio of the two quantities as a number.[39]

    a. "Without commonly agreed-upon units, it would not be possible to accurately quantify the passing of time, the length of an object, or the temperature of one's surroundings…Units allow us to count things in a building-block type fashion so they have meaning beyond a simple descriptive comparison such as smaller than, brighter than, longer than, and so on. Determination of measurement units that are deemed susceptible and repeatable, and

[34] John Quincy Adams, *Extract from the Report on Weights and Measures by the Secretary of State*, made to the Senate on February 22, 1821.
[35] JCGM, *International Vocabulary of Metrology — Basic and General Concepts and Associated Terms (VIM)*, § 2.1 (2008).
[36] ASTM, *Standard Terminology Relating to Quality and Statistics*, §3 (2009); ISO, *Medical laboratories – Particular requirements for quality and competence*, ISO 15189 §3.4 (2007); Dybkaer, *Metrology and protometrology: the ordinal question*, 12 ACCRED. QUAL. ASSUR. 553 (2007).
[37] ASTM, *Standard Guide for Defining the Test Result of a Test Method* (2003); Fuentes-Arderiu, *Vocabulary of terms in protometrology*, 11 ACCRED. QUAL. ASSUR. 640 (2006); Dybkaer, *Metrology and protometrology: the ordinal question*, 12 ACCRED. QUAL. ASSUR. 553 (2007); Krebs, *Measurement Theory*, 67(12) PHYS. THERAPY 1834, 1835 (1987).
[38] Dybkaer, *Metrology and protometrology: the ordinal question*, 12 ACCRED. QUAL. ASSUR. 553 (2007); Fuentes-Arderiu, *Vocabulary of terms in protometrology*, 11 ACCRED. QUAL. ASSUR. 640, 642 (2006).
[39] JCGM, *International Vocabulary of Metrology — Basic and General Concepts and Associated Terms (VIM)*, § 1.9 (2008).

maintaining them as measurement standards, lies at the heart of fundamental metrology concepts and principles."[40]

b. THE INTERNATIONAL SYSTEM OF UNITS (SI): The SI was established by and is defined by the General Conference on Weights and Measures in 1960. The base quantities used in the SI are length, mass, time, electric current, thermodynamic temperature, amount of substance, and luminous intensity. The corresponding base units of the SI were chosen to be the metre, the kilogram, the second, the ampere, the kelvin, the mole, and the candela.[41]

2. MEASUREMENT STANDARD: Realization of the definition of a given quantity, with stated quantity value and associated measurement uncertainty, used as a reference.[42]

a. REFERENCE MATERIAL: Object, material or substance sufficiently homogeneous and stable with reference to specified properties, which has been established to be fit for its intended use in measurement or in examination of nominal properties.[43]

i. "The use of reference materials makes possible the transfer of the values of measured or assigned quantities between testing, analytical and measurement laboratories."[44]

ii. "One of the key factors affecting laboratories' capabilities to produce reliable test data is the availability of reference materials with property values that can be relied upon by their users."[45]

iii. "A reference material is for use in a decision process, hence the requirement of reliability of the value of the property measured must be consistent with the risk associated with a wrong decision."[46]

b. REFERENCE PROCEDURE: Measurement procedure accepted as providing measurement results fit for their intended use in assessing measurement trueness of measured quantity values obtained from other measurement procedures for quantities of the same kind, in calibration, or in characterizing reference materials.[47]

[40] *The Metrology Handbook* 149 (Bucher Ed. – 2004).
[41] JCGM, *The International System of Units (SI)* §1.2 (8th ed. 2008).
[42] JCGM, *International Vocabulary of Metrology — Basic and General Concepts and Associated Terms (VIM)*, § 5.1 (2008).
[43] JCGM, *International Vocabulary of Metrology — Basic and General Concepts and Associated Terms (VIM)*, § 5.13 (2008); ISO, *Reference Materials – General and Statistical Principles for Certification*, ISO Guide 35, 2 (2006); ASTM, *Standard Terminology Relating to Forensic Science*, §4 E 1732 (2005); NIST, *Definitions of Terms and Modes Used at NIST for Value-Assignment of Reference Materials for Chemical Measurements*, NIST SP260-136, 10 (2000); NIST, *Standard Reference Materials: Handbook for SRM Users*, NIST SP260-100, 53 (1993).
[44] ISO, *General Requirements for the Competence of Reference Material Producers*, ISO Guide 34 v (2000).
[45] ILAC, *Guidelines for the Requirements for the Competence of Reference material Producers*, ILAC G12, 4 (2000); Zschunke, *The Role of Reference Materials in Analytical Chemistry*, 8 ACCRED. QUAL. ASSUR. 247, 249 (2003).
[46] NIST, *Standard Reference Materials: Handbook for SRM Users*, NIST SP260-100, 16 (1993).
[47] JCGM, *International Vocabulary of Metrology — Basic and General Concepts and Associated Terms (VIM)*, § 2.7 (2008).

8105 NE 140th Pl., Bothell WA 98011

C. MEASUREMENT AND TESTING PROCESS: The objective of a measurement is to determine the value of the particular quantity being measured.

1. SUBJECT:

a. MEASURAND: Quantity intended to be measured.[48]

i. "The specification of a measurand requires knowledge of the kind of quantity, description of the state of the phenomenon, body, or substance carrying the quantity, including any relevant component, and the chemical entities involved."[49]

ii. "The measurement, including the measuring system and the conditions under which the measurement is carried out, might change the phenomenon, body, or substance such that the quantity being measured may differ from the measurand as defined. In this case, adequate correction is necessary."[50]

b. OBSERVAND/PROTO-MEASURAND: "Particular nominal or ordinal property intended to be observed."[51]

i. EX. "In chemistry, 'analyte', or the name of a substance or compound, are terms sometimes used for 'measurand'. This usage is erroneous because these terms do not refer to quantities."[52]

2. PROCESS:

a. TEST METHOD: Defined technical procedure to determine one or more specified characteristics of a material or product.[53]

i. "Understanding the mechanics and theory behind…methods is helpful not only for determining the best method for a particular situation or application but also for understanding their limitations and the…data they provide."[54]

b. Direct Measurement: A measurement that senses the quantity of interest itself and maps it to a quantity value without the necessity of intermediate determinations.

c. Indirect Measurement: The determination of a quantity of interest through its relationship to other directly measured quantities.

[48] JCGM, *International Vocabulary of Metrology — Basic and General Concepts and Associated Terms (VIM)*, § 2.3 (2008).
[49] JCGM, *International Vocabulary of Metrology — Basic and General Concepts and Associated Terms (VIM)*, § 2.3 Note 1 (2008).
[50] JCGM, *International Vocabulary of Metrology — Basic and General Concepts and Associated Terms (VIM)*, § 2.3 Note 2 (2008).
[51] Fuentes-Arderiu, *Vocabulary of terms in protometrology*, 11 ACCRED. QUAL. ASSUR. 640, 642 (2006).
[52] JCGM, *International Vocabulary of Metrology — Basic and General Concepts and Associated Terms (VIM)*, § 2.3 Note 4 (2008).
[53] NIST, Handbook 150 § 1.5.29 (2006).
[54] *The Metrology Handbook* 157 (Bucher Ed. – 2004).

d. MEASUREMENT PROCEDURE (SOP): "Detailed description of a measurement according to one or more measurement principles and to a given measurement method, based on a measurement model and including any calculation to obtain a measurement result."[55]

e. MEASURING SYSTEM: Set of one or more measuring instruments and often other devices, including any reagent and supply, assembled and adapted to give information used to generate measured quantity values within specified intervals for quantities of specified kinds.[56]

i. "The makeup of a measurement system is determined by an application or particular situation. The adequacy of a measurement system depends on the accuracy and reliability requirements of the measurement data…How measurement data will be used will drive the selection, composition and sophistication of a measurement system in order to meet measurement objectives…For a measurement system to be properly constructed, a comprehensive understanding of applicable measurement application(s) is required…Measurement systems produce data within a window normally associated with a probability or likelihood that the data obtained faithfully represent their intended measurand(s)."[57]

3. VALIDITY: "*Validity* is the extent to which an item actually measures what the researcher purports the item measures. Measurement validity is the paramount goal of data collection."[58]

a. VALIDATION: "Validation is the confirmation by examination and the provision of objective evidence that the particular requirements for a specific intended use are fulfilled."[59]

i. "One particular task of science is the validation of new methods to determine their reliability under different conditions and their limitations."[60]

ii. ISO 17025 "includes a well established list of techniques that can be used, alone or in combination, to validate a method."[61]

iii. "The laboratory shall validate non-standard methods, laboratory-designed/developed methods, standard methods used outside their intended scope, and amplifications and modifications of standard methods to confirm that the methods are fit for the intended use…The laboratory shall record the results obtained, the procedure used for the validation, and a statement as to whether the method is fit for the intended use."[62]

---

[55] JCGM, *International Vocabulary of Metrology — Basic and General Concepts and Associated Terms (VIM),* § 2.6 (2008).
[56] JCGM, *International Vocabulary of Metrology — Basic and General Concepts and Associated Terms (VIM),* § 3.2 (2008).
[57] *The Metrology Handbook* 159-161 (Bucher Ed. – 2004).
[58] Krebs, *Measurement Theory*, 67(12) PHYS. THERAPY 1834, 1838 (1987).
[59] ISO, *General requirements for the competence of testing and calibration laboratories,* ISO 17025 § 5.4.5.1 (2005); NIST, Handbook 150 § 5.4.5.1 (2006).
[60] NAS, *Strengthening Forensic Science in the United States: A Path Forward*, 113 (2009).
[61] NAS, *Strengthening Forensic Science in the United States: A Path Forward*, 113-114 (2009).
[62] ISO, *General requirements for the competence of testing and calibration laboratories,* ISO 17025 § 5.4.5.2 (2005); NIST, Handbook 150 § 5.4.5.1 (2006).

a) Both quantitative and "[q]ualitative methods should be subjected to validation processes in order to ensure their particular fitness for purpose."[63]

b) "Validation includes specification of the requirements, determination of the characteristics of the methods, a check that the requirements can be fulfilled by using the method, and a statement on the validity."[64]

i. Validation of quantitative test methods must include statements of the uncertainty of the method such as documentation of precision and bias.[65]

ii. "The most common, and probably the most useful, form of data treatment in method-validation studies for qualitative tests is the calculation and reporting of either specificity and sensitivity or false positive and negative error rates."[66]

i. "Validation…establishes the crucial link between the metrological approach (analytical properties) and solving analytical problems (fitness for purpose)."[67]

ii. Peer review: "A critical step in such validation studies is their publication in peer reviewed journals, so that experts in the field can review, question, and check the repeatability of the results. These publications must include clear statements of the hypotheses under study, as well as sufficient details about the experiments, the resulting data, and the data analysis so that the studies can be replicated. Replication will expose not only additional sources of variability but also further aspects of the process, leading to greater understanding and scientific knowledge that can be used to improve the method."[68]

i. Computer Use:

c) "When computers or automated equipment are used for the acquisition, processing, recording, reporting, storage or retrieval of test or calibration data, the laboratory shall ensure that computer software developed by the user is documented in sufficient detail and is suitably validated as being adequate for use."[69]

i. "Commercial off-the-shelf software (e.g. word processing, database and statistical programs) in general use within their designed application range may

[63] Rios, *Quality assurance of qualitative analysis in the framework of the European project 'MEQUALAN'*, 8 ACCRED. QUAL. ASSUR. 68, 74 (2003); Rios, *Reliability of binary analytical responses*, 24(6) TRENDS ANAL. CHEM. 509, 513 (2005).
[64] ISO, *General requirements for the competence of testing and calibration laboratories*, ISO 17025 § 5.4.5.3 Note 1 (2005); NIST, Handbook 150 § 5.4.5.3 Note 1 (2006).
[65] ASTM, *Standard Guide for Statistical Procedures to Use in Developing and Applying Test Methods*, E 1488 § 4.1 (2008); ISO, *General requirements for the competence of testing and calibration laboratories*, ISO 17025 § 5.4.5.2-5.4.5.3 (2005); NIST, Handbook 150 § 5.4.5.2-5.4.5.3 (2006).
[66] Ellison, *Characterizing the performance of qualitative analytical methods: Statistics and terminology*, 24(6) TRENDS ANAL. CHEM. 468, 470 (2005).
[67] Rios, *Quality assurance of qualitative analysis in the framework of the European project 'MEQUALAN'*, 8 ACCRED. QUAL. ASSUR. 68, 74 (2003).
[68] NAS, *Strengthening Forensic Science in the United States: A Path Forward*, 114 (2009).
[69] ISO, *General requirements for the competence of testing and calibration laboratories*, ISO 17025 § 5.4.7.2 (2005); NIST, Handbook 150 § 5.4.7.2 (2006).

be considered to be sufficiently validated. However, laboratory software configuration/modifications should be validated."[70]

b. FITNESS FOR PURPOSE:

i. "Measurement results are the product of a process and not simply an instrument. Confidence in results can occur only after showing the entire program is 'fit-for-purpose.'"[71]

ii. "For an analytical result to be fit for its intended purpose it must be sufficiently reliable that any decision based on it can be taken with confidence. Thus the method performance must be validated and the uncertainty on the result, at a given level of confidence, estimated."[72]

iii. "It is generally acknowledged that the fitness for purpose of an analytical result cannot be assessed without an estimate of the measurement uncertainty to compare with the level of confidence required."[73]

D. QUALITY ASSURANCE:

1. TRACEABILITY: Property of a measurement result whereby the result can be related to a reference through a documented unbroken chain of calibrations, each contributing to the measurement uncertainty.[74]

a. "Metrological traceability is established via an identified calibration hierarchy from the stated reference to the calibrator of the final measurement. Each calibrator in the chain has its quantity value established by comparison to the preceding calibrator."[75]

b. Traceability includes the following essential elements:[76]

i. "*Unbroken chain of comparisons.* A documented system of comparisons going back to a standard acceptable to the parties, usually a national or international standard;"

ii. "*Measurement uncertainty.* The measurement uncertainty for each step in the traceability chain must be calculated according to defined methods and must be stated so that an overall uncertainty for the whole chain may be calculated;"

[70] ISO, *General requirements for the competence of testing and calibration laboratories*, ISO 17025 § 5.4.7.2 (2005); NIST, Handbook 150 § 5.4.7.2 (2006).
[71] Gullberg, *Methodology and Quality Assurance in Forensic Breath Alcohol Analysis*, 12 For. Sci. Rev. 49, 49 (2000).
[72] EURACHEM, *The Fitness for Purpose of Analytical Methods A Laboratory Guide to Method Validation and Related Topics* § 4.4 (1998).
[73] Shegunova, *Estimation of measurement uncertainty in organic analysis two practical approaches*, ACCRED. QUAL. ASSUR. (2008)
[74] JCGM, *International Vocabulary of Metrology — Basic and General Concepts and Associated Terms (VIM)*, § 2.41 (2008); NIST, Handbook 150 § 1.5.30 (2006).
[75] Hibbert, *Metrological traceability: I make it 42; you make it 42; but is it the same 42?* 13 ACCRED. QUAL. ASSUR. 11 (2006).
[76] NIST, *Good Measurement Practice for Ensuring Traceability*, GMP-13, § 1.2 (2003).

iii. "*Documentation.* Each step in the chain must be performed according to documented and generally acknowledged procedures (see GMP 12) and the results must be documented."

c. Property of a measurement result:

i. "Traceability applies to the measured value and it's uncertainty, as a single entity. One without the other is not traceable."[77]

ii. "[M]etrological traceability is a property of a measurement result…metrological traceability tells us about a measurement result, not a method, not an institute, nor a laboratory…Incorrect thinking about the 'traceability of a method' leads to the implication that the analytical system will somehow always be traceable. Unfortunately, this is not correct; every measurement that is made must be shown to be traceable."[78]

d. Comparability of Measurement Results:

i. "Laboratory tests are usually performed to assist the person requesting the test to make a decision. The result of a test is often compared to a limit, reference interval or another test result obtained previously. Meaningful comparisons can only be made if results are traceable to a common reference and the uncertainty of measurement relative to that common reference is known."[79]

ii. "Traceability provides the terminology, concepts and strategy for ensuring that…measurements are comparable…Traceability is a concept and a measurement strategy which provides a means of anchoring measurements in both time and space…Measurements made at different times or in different places are directly related to a common reference."[80]

iii. "Comparability is an essential property of analytical results."[81]

e. Accuracy and Reliability:

i. "Traceability ensures that the measurements are accurate representations of the specific quantity subject to measurement, within the uncertainty of the measurement."[82]

---

[77] *The Metrology Handbook* 65 (Bucher Ed. – 2004).
[78] Hibbert, *Metrological traceability: I make it 42; you make it 42; but is it the same 42?* 11 ACCRED. QUAL. ASSUR. 543, 545 (2006).
[79] *Uncertainty Of Measurement In Biological, Forensic, Medical And Veterinary Testing*, NATA TECH. CIRC. 1 (December 2003).
[80] King, *Perspective: Traceability of Chemical Analysis*, 122 ANALYST 197, 197 (1997); Hibbert, *Metrological traceability: I make it 42; you make it 42; but is it the same 42?* 11 ACCRED. QUAL. ASSUR. 543, 546 (2006); ISO, *Reference Materials – General and Statistical Principles for Certification*, ISO Guide 35, § 1 (2006).
[81] Ellison, *Using validation data for ISO measurement uncertainty estimation Part 1. Principles of an approach using cause and effect analysis*, 123 ANALYST 1387 (1998).
[82] NIST, *Good Measurement Practice for Ensuring Traceability*, GMP-13, § 1.1 (2003).

8105 NE 140th Pl., Bothell WA 98011

ii. "Among the many aspects of measurement that affect reliability, metrological traceability is essential. It underpins the ability of the analyst to claim that his or her result is what it purports to be."[83]

iii. "The measurement of known and traceable standards is the basis for determining accuracy and thereby confidence in all analytical results."[84]

iv. "It is not possible to determine a reliable result and its uncertainty if there is no traceability of the measurement to a standard with known uncertainty."[85]

2. CALIBRATION: Operation that, under specified conditions, in a first step, establishes a relation between the quantity values with measurement uncertainties provided by measurement standards and corresponding indications with associated measurement uncertainties and, in a second step, uses this information to establish a relation for obtaining a measurement result from an indication.[86]

a. When required:

i. "All equipment used for tests and/or calibrations…having a significant effect on the accuracy or validity of the result of the test, calibration or sampling shall be calibrated before being put into service. The laboratory shall have an established programme and procedure for the calibration of its equipment."[87]

ii. "Any instrument or artifact used as part of the measurement process must recently have been calibrated by reference to a standard that is traceable to a primary standard."[88]

iii. "Measurement processes are dynamic systems and often deteriorate with time or use…A calibration performed only once establishes a one-time reference of uncertainty. Recalibration detects uncertainty growth and serves to reset values while keeping a bound on the limits of errors. A properly selected interval assures that an item will receive recalibration at the proper time."[89]

b. Uncertainty in Calibration:

i. Despite its importance, all "calibration…involves uncertainty."[90]

[83] IUPAC, *Metrological Traceability of Measurement Results in Chemistry*, DRAFT § 1.2 (2008).
[84] Gullberg, *Using a Weighted Mean to Compute the Values of Simulator Solution Standards*, 14(3) J. ANAL. TOXICOL. 196 (1990).
[85] Knopf, *Traceability system for breath-alcohol measurements in Germany*, XLVII(2) OIML BULL. 15, 17 (2007).
[86] JCGM, *International Vocabulary of Metrology — Basic and General Concepts and Associated Terms (VIM)*, § 2.39 (2008).
[87] ISO, *General requirements for the competence of testing and calibration laboratories*, ISO 17025 § 5.6.1 (2005).
[88] Kirkup, *An Introduction to Uncertainty in Measurement* 31 (Cambridge University Press 2006).
[89] NIST, *Good Laboratory Practice for Assignment and Adjustment of Calibration Intervals for Laboratory Standards*, GLP-11, 1 (2003).
[90] Bich, *Interdependence between measurement uncertainty and metrological traceability* ACCRED. QUAL. ASSUR. (IN PRESS - 2009).

8105 NE 140th Pl., Bothell WA 98011

ii. "A calibration is not complete until the expanded uncertainty associated with the calibration is determined and reported."[91]

iii. "The uncertainty of the calibration will depend on the uncertainty of the values of the standards and the measurement processes used for the intercomparisons."[92]

iv. "[U]se of proper standards and equipment, and selection of standard operating procedures are essential for providing calibration results with accurate and traceable values with appropriate and suitable uncertainties."[93]

c. Calibration defines the valid range of measurement:

i. "Standards should never be used in an extrapolative mode. They should always bracket the measurement range. No measurement should be reported at a value lower or higher than the lowest or highest standard used to calibrate the measurement process."[94]

ii. "It is not good measurement practice to report extrapolated data, i.e., outside the range calibrated."[95]

iii. "It is a generally accepted principle of reliable analysis that chemical analyzers should be calibrated over the full range of measurement and that measurement data be restricted to the range calibrated."[96]

d. "Calibration with proper standards is the key to metrological traceability."[97]

3. Traceability and Calibration in Qualitative Test Observations:

a. Reference materials and procedures "are the key elements in assuring traceability of the qualitative results/information."[98]

b. "Traceability of measurement results, reference values and calibration values is essential in qualitative testing. It is particularly critical where the qualitative test relies on comparison with reference values."[99]

4. QUALITY ASSURANCE PROGRAM: "The laboratory shall have quality control procedures for monitoring the validity of tests and calibrations undertaken. The resulting data shall be recorded

[91] NIST, *Good Laboratory Practice for Rounding Expanded Uncertainties and Calibration Values*, GLP-9, 1 (2003).
[92] NIST, *Standard Reference Materials: Handbook for SRM Users*, NISTSP 260-100, 6 (1993).
[93] NIST, *Good Measurement Practice for Standard Operating Procedure Selection*, GMP-12, 1 (2003).
[94] NIST, *Standard Reference Materials: Handbook for SRM Users*, NISTSP 260-100, 6 (1993).
[95] NIST, *Standard Reference Materials: Handbook for SRM Users*, NISTSP 260-100, 7 (1993).
[96] NIST, *Standard Reference Materials: Handbook for SRM Users*, NISTSP 260-100, 7 (1993).
[97] Hibbert, *Metrological traceability: I make it 42; you make it 42; but is it the same 42?* 11 ACCRED. QUAL. ASSUR. 543, 543 (2006).
[98] Rios, *Reliability of binary analytical responses*, 24(6) TRENDS ANAL. CHEM. 509, 510 (2005).
[99] Ellison, *Uncertainties in qualitative testing and analysis*, 5 ACCRED. QUAL. ASSUR. 346, 348 (2000).

in such a way that trends are detectable and, where practicable, statistical techniques shall be applied to the reviewing of the results."[100]

a. ACCREDITATION: An independent authoritative body gives formal recognition that a lab adheres to an established set of standards of quality and relies on acceptable practices within these requirements to render it competent to carry out specific tests or calibrations or types of tests or calibrations.[101]

   i. "Accrediting bodies require that the methods meet a level of acceptable practice."[102]

      a) "Laboratories shall be able to demonstrate proper use of traceable standards and test and measurement equipment by competent laboratory personnel in a suitable environment in performing the tests for which accreditation is desired or held. This demonstration will include the determination of the appropriate measurement uncertainty."[103]

   ii. Best Measurement Capability: "Smallest uncertainty of measurement a laboratory can achieve within its scope of accreditation, when performing more-or-less routine calibrations of nearly ideal measurement standards intended to define, realize, conserve or reproduce a unit of that quantity or one or more of its values, or when performing more-or-less routine calibrations of nearly ideal measurement instruments designed for the measurement of that quantity."[104]

   iii. Scope of Accreditation: "The Scope of Accreditation lists the test methods or services, or calibration services, for which the laboratory is accredited."[105]

b. PROFICIENCY TESTING: Determination of laboratory testing performance by means of interlaboratory comparisons.[106]

   i. "Proficiency testing requirements are associated with most fields of accreditation."[107]

   ii. "The performance of tests or calibrations and reporting of results from proficiency testing assists…in determining a laboratory's competence and the effectiveness of its management system. Information obtained from proficiency testing helps to identify technical problems in a laboratory."[108] Types of processes subject to proficiency testing include:[109]

[100] ISO, *General requirements for the competence of testing and calibration laboratories*, ISO 17025 § 5.9.1 (2005); NIST, Handbook 150 § 5.9.1 (2006).
[101] NIST, Handbook 150 § 1.5.1 (2006); NAS, *Strengthening Forensic Science in the United States: A Path Forward*, 7-2 (2009).
[102] NAS, *Strengthening Forensic Science in the United States: A Path Forward*, 7-10 (2009).
[103] NIST, Handbook 150 App. B.2 (2006).
[104] NIST, Handbook 150 § 1.5.5 (2006).
[105] NIST, Handbook 150 § 1.5.26 (2006).
[106] NIST, Handbook 150 § 1.5.21 (2006).
[107] NIST, Handbook 150 § 3.4.2.1 (2006).
[108] NIST, Handbook 150 § 3.4.1.1 (2006).
[109] NIST, Handbook 150 § 1.5.21 (2006).

a) Sampling—for example, where individuals or organizations are required to take samples for subsequent analysis;

b) Qualitative schemes—for example, where laboratories are required to identify a component of a test item; and

c) Data transformation—for example, where laboratories are furnished with sets of data and are required to manipulate the data to provide further information.

iii. "Proficiency testing has long been recognized among analytical chemists as useful for evaluating instrumental, method, laboratory and program performance."[110]

E. MEASUREMENT INTERPRETATION:

1. "It is scientific only to say what is more likely and what is less likely."[111]

a. "Even when an analytical procedure has been performed correctly and precisely, variables can affect the test result. Knowledge of these variables and standardization of laboratory testing procedures are essential for correct interpretation and optimal use of the data."[112]

2. Measurement Result: Set of quantity values being attributed to a measurand together with any other available relevant information.[113]

a. The value of a measurand can never be known exactly; all that can be known is its estimated value.[114]

i. "Every measurement has an uncertainty associated with it, resulting from errors arising in the various stages of sampling and analysis and from imperfect knowledge of factors affecting the result. For measurements to be of practical value it is necessary to have some knowledge of their reliability or uncertainty."[115]

a) Ex. We wish to know the quantity Y associated with a substance being measured. Given that the exact value of Y can never be known, we chose to make multiple measurements and average them to arrive at a best estimate. Our best estimate can be expressed as:[116]

$$Y = \bar{y} + \varepsilon$$

where
$\bar{y}$ = mean of measurements

[110] Gullberg, *Results of a Proposed Breath Alcohol Proficiency Test Program*, 51(1) J. For. Sci. 168,168 (2006).
[111] Feynman, *The Character of Physical Law* 165-166 (MIT Press 1965).
[112] NCCLS, Procedures for the Collection of Diagnostic Blood Specimens by Venipuncture; Approved Standard—Fifth Edition, H3-A5, § 5 (2003).
[113] JCGM, *International Vocabulary of Metrology — Basic and General Concepts and Associated Terms (VIM)*, § 2.9 (2008).
[114] JCGM, *Evaluation of measurement data — Guide to the expression of uncertainty in measurement (GUM)*, Appendix D.4 (2008); Kirkup, *An Introduction to Uncertainty in Measurement* 33 (Cambridge University Press 2006).
[115] EURACHEM, *Guide to Quality in Analytical Chemistry* § 16.1 (2002).
[116] Eleftheriou, *Measuring performance in analytical measurements* 14 ACCRED. QUAL. ASSUR. 67, 67 (2009).

ε = unknown uncertainty associated with mean

b. UNCERTAINTY: "Characterization of the dispersion of values assignable to a measurand based on the information available including systematic and random effects, definitional uncertainty and any other factors that may impact the measurement or test process or result."[117]

i. Uncertainty is a property of quantitative measurement results.[118]

ii. The estimate of uncertainty of a measurement:

a) "Quantifies the quality of a measurement result."[119]

b) "Reflects the lack of exact knowledge of the value of the measurand."[120]

c) "Is a necessary step in producing traceable results."[121]

iii. "Knowledge of the uncertainty associated with measurement results is essential to the interpretation of the results. Without quantitative assessments of uncertainty, it is impossible to decide whether observed differences between results reflect more than experimental variability, whether test items comply with specifications, or whether laws based on limits have been broken. Without information on uncertainty, there is a risk of misinterpretation of results. Incorrect decisions taken on such a basis may result in unnecessary expenditure in industry, incorrect prosecution in law, or adverse health or social consequences."[122]

3. OBSERVATION RESULT: "[E]stimated value of a particular nominal or ordinal property, obtained by observation."[123]

a. "Qualitative analysis is characterized by its binary nature: presence/absence, positive sample/negative sample, or yes/no according to a pre-set threshold." [124]

i. Types of qualitative analysis:[125]

[117] JCGM, *International Vocabulary of Metrology — Basic and General Concepts and Associated Terms (VIM)*, § 2.26 (2008); ASTM, *Standard Terminology Relating to Quality and Statistics*, E 456 § 3 (2008).
[118] Rios, *Quality assurance of qualitative analysis in the framework of the European project 'MEQUALAN'*, 8 ACCRED. QUAL. ASSUR. 68, 71 (2003).
[119] Croarkin, *Statistics and Measurements* 106 J. RES. NATL. INST. STAND. TECHNOL. 279, 283 (2001).
[120] JCGM, *Evaluation of measurement data — Guide to the expression of uncertainty in measurement (GUM)*, § 3.3.1 (2008).
[121] *Uncertainty Of Measurement In Biological, Forensic, Medical And Veterinary Testing*, NATA TECH. CIRC. 1 (December 2003).
[122] ISO, *Guidance for the use of repeatability, reproducibility and trueness estimates in measurement uncertainty estimation*, ISO/TS 21748 DRAFT REVISION, v (2009).
[123] Fuentes-Arderiu, *Vocabulary of terms in protometrology*, 11 ACCRED. QUAL. ASSUR. 640, 642 (2006).
[124] Rios, *Quality assurance of qualitative analysis in the framework of the European project 'MEQUALAN'*, 8 ACCRED. QUAL. ASSUR. 68, 69 (2003); Rios, *Reliability of binary analytical responses*, 24(6) TRENDS ANAL. CHEM. 509, 512 (2005).
[125] Rios, *Quality assurance of qualitative analysis in the framework of the European project 'MEQUALAN'*, 8 ACCRED. QUAL. ASSUR. 68, 69 (2003).

a) Identification.

b) Classification.

ii. "It is important to recognize…that any method or technique used for classification purposes, no matter how simple it may be to perform, will eventually fail to classify all samples correctly."[126] This is true "even when the analyst making the identification follows all the canons of best practice."[127]

iii. "Interpretation of the results must accordingly take the relevant uncertainties into account."[128]

iv. Uncertainty in qualitative methods is generally associated with the probabilistic determination of the reliability/unreliability of a method.[129]

b. UNRELIABILITY: The unreliability of a qualitative method is a measure of its likelihood of giving an erroneous response (error rate).[130]

i. "Traceability and (un)reliability of the [results] produced by these methods are crucial parameters in assuring the quality expected of the information derived."[131]

4. STATISTICAL CONCEPTS:

a. Population: The entire set or universe of objects sharing specific traits defining a class of objects.

b. Sample: A subset of objects selected from the population.

c. Distribution: The set of possible values of a random variable related through their frequency of occurrence or belief based relative likelihood.

d. Parameter: A characteristic of a population's distribution.

e. Statistic: A characteristic of a sample's distribution.

[126] Lendl, *Advancing from unsupervised, single variable-based to supervised, multivariate-based methods: A challenge for qualitative analysis*, 24(6) TRENDS ANAL. CHEM. 488, 488 (2005).
[127] Ellison, *Quantifying uncertainty in qualitative analysis* 123 ANALYST 1155, 1155 (1998).
[128] Ellison, *Quantifying uncertainty in qualitative analysis* 123 ANALYST 1155, 1155 (1998).
[129] Rios, *Quality assurance of qualitative analysis in the framework of the European project 'MEQUALAN'*, 8 ACCRED. QUAL. ASSUR. 68, 71 (2003); Mil'man, *Uncertainty of Qualitative Chemical Analysis: General Methodology and Binary Test Methods*, 59(12) J. ANAL. CHEM. 1128, 1130-1134, 1136 (2004); Ellison, *Characterizing the performance of qualitative analytical methods: Statistics and terminology*, 24(6) TRENDS ANAL. CHEM. 468, 469-70 (2005); Lewis, *Reliability and Validity: Meaning and Measurement*, 10-11, Presentation to Annual Meeting of the Society for Academic Emergency Medicine (1999); ISO, *Statistics — Vocabulary and symbols — Part I: General statistical terms and terms used in probability*, ISO 3534-1 §§ 1.46, 1.47 (2006).
[130] Mil'man, *Uncertainty of Qualitative Chemical Analysis: General Methodology and Binary Test Methods*, 59(12) J. ANAL. CHEM. 1128, 1128 (2004); Ellison, *Uncertainties in qualitative testing and analysis*, 5 ACCRED. QUAL. ASSUR. 346, 347 (2000); Rios, *Quality assurance of qualitative analysis in the framework of the European project 'MEQUALAN'*, 8 ACCRED. QUAL. ASSUR. 68, 70-74 (2003).
[131] Rios, *Reliability of binary analytical responses*, 24(6) TRENDS ANAL. CHEM. 509, 515 (2005).

f. Descriptive Statistics: Utilizes data to describe the properties of a sample, not to make predictions based upon it.

g. Inferential Statistics: Utilizes data to draw inference or make predictions. A typical example is the use of sample data to generate a sample statistic from which an inference concerning a population parameter may be made.

h. Probability – Frequentist Interpretation: Probability is interpreted as relative frequency of occurrence over all sample data sets. As such, probabilities are objectively determined as a function of sampling data. Population parameters have unique, fixed true values that are unknown. The randomness lies in the sampling process, not the parameter. Since population parameters are nonrandom, probability statements cannot be made about their values. Nor can probability statements be made about a characteristic of a unique event. The parameter or characteristic either is or is not a particular value. The level of confidence associated with an inference refers to the confidence in the sampling/inferential process, *not* the actual quantity of interest. It tells us how often, over repeated samplings, our inference will happen to correspond to the true value.

i. Probability – Bayesian Interpretation: Probability is interpreted as an information-based "degree of belief" that an event will occur. Bayesian inference employs sampling data and any other information deemed relevant in the decision making process so that probability (degree of belief) may be based upon both objective and subjective components. In this framework, the parameters themselves are considered random so that probability statements can be made directly about their values. The same holds for a characteristic of a unique event. Thus, probability statements made concerning the value of a parameter or characteristic *are* about the actual quantity of interest. It tells us the probability that this particular inference is "true".

j. The philosophy underlying each of these approaches is profoundly distinct.[132] The frequentist interpretation is that most widely espoused although Bayesian theory has gained prominence.[133] Regardless, both rely upon essentially the same statistical tools. Moreover, the methods are often combined and the usefulness of either approach depends upon the circumstances of the measurement, the validity of any assumptions and the use to be made of the results. One should be aware of both approaches to be able to adequately evaluate uncertainty/unreliability claims concerning a test result.

5. DETERMINATION OF UNCERTAINTY AND UNRELIABILITY:

a. "Testing laboratories shall have and shall apply procedures for estimating uncertainty of measurement. In certain cases the nature of the test method may preclude rigorous,

[132] Howson, *Scientific Reasoning The Bayesian Approach* 20-21 (Open Court 2006); Ehrlich, *Evolution of philosophy and description of measurement* 12 ACCRED. QUAL. ASSUR. 201, 205 (2007); Brüchle, *Confidence intervals for experiments with background and small numbers of events* 91 RADIOCHIM. ACTA 71, 74 (2003); D'Agostini, *Role and Meaning of Subjective Probabaility*, 568 AIP Conference Proceedings 23 (2001); Estler, *Measurement as Inference: Fundamental Ideas*, 48(2) Annals of the CIRP 611, 618 (1999); D'Agostini, *Bayesian Reasoning Versus Conventional Statistics in High Energy Physics*, presentation at XVIII International Workshop on Maximum Entropy and Bayesian Methods (Germany 1998).

[133] Croarkin, *Statistics and Measurements* 106 J. RES. NATL. INST. STAND. TECHNOL. 279, 290-291 (2001).

metrologically and statistically valid, calculation of uncertainty of measurement. In these cases the laboratory shall at least attempt to identify all the components of uncertainty and make a reasonable estimation, and shall ensure that the form of reporting of the result does not give a wrong impression of the uncertainty. Reasonable estimation shall be based on knowledge of the performance of the method and on the measurement scope and shall make use of, for example, previous experience and validation data."[134]

b. MEASUREMENTS V. OBSERVATIONS:

i. "Traditional metrological principles, as they are applied to quantitative methods, cannot be directly applied to qualitative ones."[135] Accordingly quantitative and qualitative methods are treated separately.

c. UNCERTAINTY OF MEASUREMENT RESULTS (QUANTITATIVE METHODS):

i. "The approach to quantification of uncertainty in measurement, which is now widely used in the physical sciences, is that presented in the *Guide to the Expression of Uncertainty in Measurement.*"[136]

ii. Basic Concepts:

a) APPROXIMATE MEASURAND VALUE: It is understood that the result of a measurement is merely an approximation of the quantity value attributable to the measurand. In the first step of measurement analysis, this typically consists of either an arithmetic or weighted mean. When multiple measurements are obtained, the best estimate of Y may be based on either an arithmetic or weighted mean.[137] Although in special circumstances the weighted and classical mean may be equal, in general they will not be.[138]

1) ARITHMETIC MEAN:[139] This is a simple average of measurement values. It is determined by adding all measured values together and then dividing the sum by the number of values included in the sum. It is typically used when all measured values are considered to be equally reliable.

$$\bar{y} = \frac{1}{N} \cdot \sum_{i=1}^{N} y_i$$

[134] ISO, *General requirements for the competence of testing and calibration laboratories*, ISO 17025 § 5.4.6.2 (2005).
[135] Rios, *Quality assurance of qualitative analysis in the framework of the European project 'MEQUALAN'*, 8 ACCRED. QUAL. ASSUR. 68, 69 (2003); Rios, *Reliability of binary analytical responses*, 24(6) TRENDS ANAL. CHEM. 509, 512 (2005).
[136] Toman, *Bayesian Approach to Assessing Uncertainty and Calculating a Reference Value in Key Comparison Experiments*, 110 J. RES. NATL. INST. STAND. TECHNOL. 605, 606 (2005).
[137] JCGM, *Evaluation of measurement data — Guide to the expression of uncertainty in measurement (GUM)*, § 4.1.4 (2008); NIST, *Standard Reference Materials: Handbook for SRM Users*, NISTSP 260-100, 76-78 (1993).
[138] Paule, *Consensus Values and Weighting Factors*, 87 J. RES. NAT'L BUREAU STAND. 377, 380 (1982).
[139] ISO, *Statistics — Vocabulary and symbols — Part 1: General statistical terms and terms used in probability*, ISO 3534-1, § 1.15 (2007).

2) WEIGHTED MEAN:[140] When combining multiple values determined for a given measurand, a weighted mean attaches more weight to those values considered more reliable. It is determined as the sum of measurement values that have been assigned relative weights based on the importance or confidence we have in a particular measurement divided by the sum of the weights.

$$\bar{y}_{wm} = \frac{\sum_{i=1}^{N} w_i \cdot y_i}{\sum_{i=1}^{N} w_i}$$

where
ω_i = weighting factor

3) Traditional weighted mean: Frequently, the values to be combined are the arithmetic means from several sets of measurements. The traditional weighted mean relies upon the precision associated with each set of measurements to determine the weight to accord the mean associated with each set. The greater the precision associated with a given mean, the more confidence we have in the value, and the more weight it is accorded in combining the means to determine a best estimate of the true value. In this case the above expression becomes:

$$\bar{y}_{wm} = \frac{\sum_{i=1}^{N} \frac{n_i}{\sigma_i^2} \cdot \bar{y}_i}{\sum_{i=1}^{N} \frac{n_i}{\sigma_i^2}}$$

The weighted mean should be employed when the values to be combined are not equally reliable.

4) Examples

ii. Within Laboratory Measurements: All measurements performed utilizing "the same method under the same conditions, that is, by the same operator, with the same equipment, on the same day and in a single laboratory."[141]

1) Arithmetic mean is appropriate.

iii. Between Laboratory Measurement: Some measurements performed where either method, conditions, analysts, operators, instruments or laboratories are different.[142]

[140] Taylor, *An Introduction to Error Analysis: The Study of Uncertainties in Physical Measurements*, 175-6 (2nd Ed. 1997); Kachigan, Statistical Analysis: An Interdisciplinary Introduction to Univariate & Multivariate Methods 49 (Radius Press 1986); Paule, *Consensus Values and Weighting Factors*, 87 J. RES. NAT'L BUREAU STAND. 377, 378 (1982).
[141] Jones, *Dealing with Uncertainty in Chemical Measurements*, 14(1) NEWSL. OF THE INT. ASSOC. FOR CHEM. TEST. 8 (2003).
[142] Jones, *Dealing with Uncertainty in Chemical Measurements*, 14(1) NEWSL. OF THE INT. ASSOC. FOR CHEM. TEST. 8 (2003); Zhang, *The Uncertainty Associated with the Weighted Mean of Measurement Data*, 43 METROLOGIA 195, 195 (2006).

1) Weighted mean accepted approach.[143]

a) Ex.: Precision of between laboratory measurements different. When the precision between sets of measurements is significant, the weighted mean should be utilized and we may employ the following weighting factor: [144]

$$w_i = \frac{n_i}{\sigma_i^2}$$

b) In this context, the weighting factor gives greater weight to those measurement results that are more precise, coinciding with the greater level of confidence in those results.[145]

iv. Under the principle of maximum likelihood, the weighted mean yields the most precise value for the best estimate of Y.[146] Failure to utilize the weighted mean in these circumstances can result in an underestimation of uncertainty.[147] "There are many situations in which it would be very misleading to average quantities without [weighting them]".[148]

b) ERROR ANALYSIS: Traditionally, the quality of a measurement result was addressed through error analysis. This approach considered each measurand as having a unique true value. "The objective of measurement in the Error Approach is to determine an estimate of the true value that is as close as possible to that single true value. The deviation from the true value is composed of random and systematic errors."[149]

i. ACCURACY: The degree of agreement of a measured value with the "true" value of the quantity of interest. A measurement is said to be more accurate when it offers a smaller measurement error. The degree of agreement expected from a

[143] Paule, *Consensus Values and Weighting Factors*, 87 J. RES. NAT'L BUREAU STAND. 377, 380 (1982); Taylor, *An Introduction to Error Analysis: The Study of Uncertainties in Physical Measurements*, 175-6 (2nd Ed. 1997); Zhang, *The Uncertainty Associated with the Weighted Mean of Measurement Data*, 43 METROLOGIA 195, 195 (2006); NIST, *Standard Reference Materials, Statistical Aspects of the Certification of Chemical Batch SRMs*, NIST SP260-125 § 8 (1996); NIST, *Standard Reference Materials: Handbook for SRM Users*, NIST SP260-100, 78 (1993); ISO, *Reference Materials – General and Statistical Principles for Certification*, ISO Guide 35, App. B.7 (2006).

[144] Zhang, *The Uncertainty Associated with the Weighted Mean of Measurement Data*, 43 METROLOGIA 195, 195 (2006); Dimech, *Calculating Uncertainty of Measurement for Serology Assays by Use of Precision and Bias* 52(3) CLIN. CHEM. 526, 527 (2006); Paule, *Consensus Values and Weighting Factors*, 87 J. RES. NAT'L BUREAU STAND. 377, 380 (1982); NIST, *Standard Reference Materials: Handbook for SRM Users*, NISTSP 260-100, 78 (1993); Witkovsky, *On Statistical Models for Consensus Values* 1(1) MEAS. SCI. REV. 33, 35 (2001).

[145] Dieck, *Measurement Uncertainty Methods and Applications* 154-155 (4th ed. 2007).

[146] Taylor, *An Introduction to Error Analysis: The Study of Uncertainties in Physical Measurements*, 175-6 (2nd Ed. 1997); Bevington, *Data Reduction and Error Analysis for the Physical Sciences* 57 (3rd 2003).

[147] Dieck, *Measurement Uncertainty Methods and Applications* 155 (4th ed. 2007); Zhang, *The Uncertainty Associated with the Weighted Mean of Measurement Data*, 43 METROLOGIA 195 (2006).

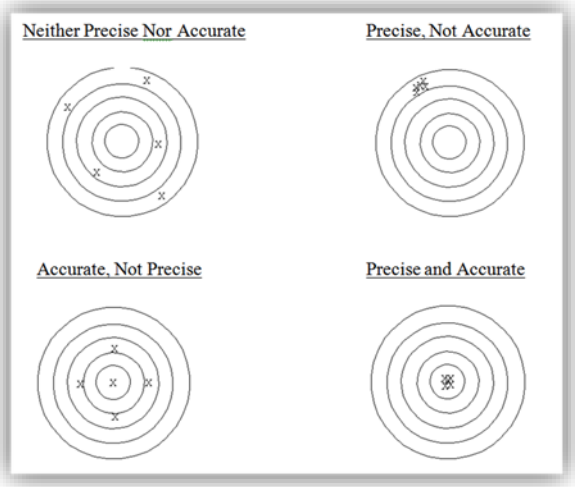
[148] Freund, *Modern Elementary Statistics* 39 (4th 1973).

[149] JCGM, *International Vocabulary of Metrology — Basic and General Concepts and Associated Terms (VIM)*, § 0.1 (2008).

measurement method/instrument is typically determined by comparing the mean of a set of measurements of a reference standard to the accepted value of the reference standard. Whether a measurement or instrument/method is deemed accurate is not an absolute judgment. Rather, accuracy is judged with respect to the use to be made of the data. What might be deemed accurate in one set of circumstances may not be accurate in another. Accuracy is not a quantity and is not given a numerical quantity value. [150]

ii. PRECISION: Closeness of agreement between indications or measured quantity values obtained by replicate measurements on the same or similar objects under specified conditions.[151] Precision is concerned with the variability or scatter of the individual results of replicate measurements. Measurements that are tightly grouped are considered precise while those with greater scatter are less so. As was the case with accuracy, precision is judged with respect to the use to be made of the data. What may be considered precise for one purpose may not be precise for another.

iii. A set of measurements may be neither accurate nor precise, precise but not accurate, accurate but not precise or both accurate and precise.[152]



iv. Measurement Interpretation – I: If a measurement value is to be interpretable, we must have an understanding of *how accurate* and *how precise* the measurement is. Absent such information, a measured value is simply a number, the meaning of which we know little about. Ideally, important measurements would be both accurate and precise. That is, not only would such

[150] JCGM, *International Vocabulary of Metrology — Basic and General Concepts and Associated Terms (VIM)*, § 2.13 (2008).
[151] JCGM, *International Vocabulary of Metrology — Basic and General Concepts and Associated Terms (VIM)*, § 2.15 (2008).
[152] Dimech, *Calculating Uncertainty of Measurement for Serology Assays by Use of Precision and Bias* 52(3) CLIN. CHEM. 526, 527 (2006).

measurements yield mean values in close agreement with a "true" value, but individual values having a high degree of agreement with each other.

An objective characterization of accuracy and precision are necessary in order to determine the value of the particular quantity being measured. Such objective characterization can be supplied by statistics.

c) MEASUREMENT ERROR: Measured quantity value minus a reference quantity value.[153] "Traditionally, an error is viewed as having two components, namely, a random component and a systematic component."[154]

i. SYSTEMATIC ERROR: Component of measurement error that in replicate measurements remains constant or varies in a predictable manner.[155] The tendency of a set of measurements to consistently (on average) underestimate or overestimate the "true" value of the measurand by a given value or percentage. Most measurements have some amount of systematic error associated with them. Systematic error may be related to measuring methods, instruments or even empirically based calculations. It is a primary component of accuracy as it has a direct and regular impact on the degree of agreement of a measured value with the "true" value of the quantity of interest. Accordingly, "if a systematic error has not been accounted for, all [measured] values could be misleading."[156] Fortunately, once identified systematic error can be corrected for. "It is assumed that the result of a measurement has been corrected for all recognized significant systematic effects and that every effort has been made to identify such effects."[157]

1) BIAS: Quantitative measure of systematic error. Bias is typically treated as either having a constant magnitude across a range of measured values or being proportional to the measured value obtained. When proportional, the bias is commonly reported as a percent bias. For chemical measurements, it is not uncommon for the bias to be proportional to measured values. "Whenever the true value of the measured quantity is needed…bias can be a serious problem."[158] Fortunately, once bias has been determined, systematic error can be easily accounted for. The bias of a method or instrument is ordinarily determined by comparing the mean of a set of measurements of a reference standard to its accepted value.[159]

$$b_c = \bar{y} - Y_{ref}$$

[153] JCGM, *International Vocabulary of Metrology — Basic and General Concepts and Associated Terms (VIM)*, § 2.16 (2008).
[154] JCGM, *Evaluation of measurement data — Guide to the expression of uncertainty in measurement (GUM)*, § 3.2.1 (2008).
[155] JCGM, *International Vocabulary of Metrology — Basic and General Concepts and Associated Terms (VIM)*, § 2.17 (2008).
[156] Les Kirkup, *An Introduction to Uncertainty in Measurement* 33 (Cambridge University Press 2006).
[157] BIPM, *Evaluation of measurement data — Guide to the expression of uncertainty in measurement (GUM)*, § 3.2.4 (2008); NIST, *Guidelines for Evaluating and Expressing the Uncertainty of NIST Measurement Results*, NIST TN 1297 § 5.2 (1994).
[158] NIST, *NIST Special Publication 260-100*, 4 (1993).
[159] JCGM, *International Vocabulary of Metrology — Basic and General Concepts and Associated Terms (VIM)*, § 2.18 (2008).

$$b_{\%} = \frac{\bar{y} - Y_{ref}}{Y_{ref}}$$

"Whenever the true value of the measured quantity is needed or when data from different laboratories, different methodologies or from the same laboratory using the same method over a period of time need to be interrelated, bias can be a serious problem."[160]

2) Best Estimate of True Value: The measurement mean corrected for bias. The bias corrected mean is often considered the best estimate of the "true" value of the measurand. Whenever reporting the mean of a set of measurement, it should be corrected for bias. The correction applied depends upon whether the bias is constant or proportional.

$$\bar{y}_b = \bar{y} - b_c$$

$$\bar{y}_b = \frac{\bar{y}}{1 + b_{\%}}$$

[160] NIST, *Standard Reference Materials: Handbook for SRM Users*, NISTSP 260-100, 4 (1993).

ii. RANDOM ERROR: Component of measurement error that in replicate measurements varies in an unpredictable manner.[161] The unpredictable/random fluctuation in measurement results under fixed conditions. Random error is associated with precision. Unlike systematic error, random error cannot be corrected for. It is an inherent aspect of all measurement results. Although random error cannot be completely eliminated, it can be minimized by making a large number of measurements.

1) MEASUREMENT STANDARD DEVIATION:[162] Quantitative characterization of the variability/dispersion of individually measured values about their mean. The standard deviation is the root mean square deviation of measured values from their mean. Precision/random error is typically expressed in terms of a standard deviation. The determination of the standard deviation varies slightly depending on the source of our data. If the standard deviation has been determined from a population, we use what is commonly referred to as a population standard deviation. On the other hand, when our data comes from a sample, we use what is commonly referred to as a sample standard deviation. Throughout the remainder of this section the distinction will not be noted unless necessary but it is assumed that whenever employed, the correct standard deviation is utilized.

$$\sigma_{y_p} = \sqrt{\frac{1}{N} \cdot \sum_{i=1}^{N} (\bar{y} - y_i)^2}$$

$$\sigma_{y_s} = \sqrt{\frac{1}{n-1} \cdot \sum_{i=1}^{n} (\bar{y} - y_i)^2}$$

2) Standard deviation (error) of the mean: Quantitative characterization of the variability/dispersion of sample means. Due to the Central Limit Theorem, the following relationship holds regardless of the underlying population distribution as long as the sample size is large enough.

$$\sigma_{\bar{y}} = \frac{\sigma_y}{\sqrt{N}}$$

[161] JCGM, *International Vocabulary of Metrology — Basic and General Concepts and Associated Terms (VIM)*, § 2.19 (2008).
[162] JCGM, *Evaluation of measurement data — Guide to the expression of uncertainty in measurement (GUM)*, § 4.2.2 (2008); Kirkup, *An Introduction to Uncertainty in Measurement* 57 (Cambridge University Press 2006); Dieck, *Measurement Uncertainty Methods and Applications* 46 (4th ed. 2007).

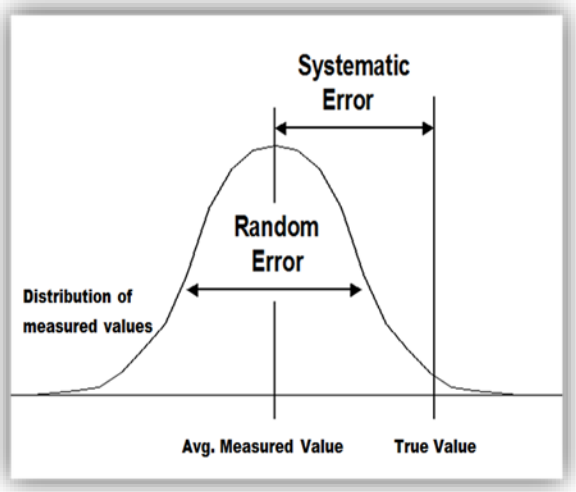
3) Standard deviation of the Traditional Weighted Mean:

$$\sigma_{wm} = \frac{1}{\sqrt{\sum \frac{n_i}{\sigma_i^2}}}$$

4) Coefficient of Variation: The standard deviation expressed as a proportion relative to the mean of a set of measurements. The coefficient of variation can be useful when combining standard deviations or comparing the variability of separate measurements.

$$\text{cv} = {}^{\sigma}/_{\bar{y}}$$

iii. "Error analysis is the attempt to estimate the total error using frequency-based statistics."[163]



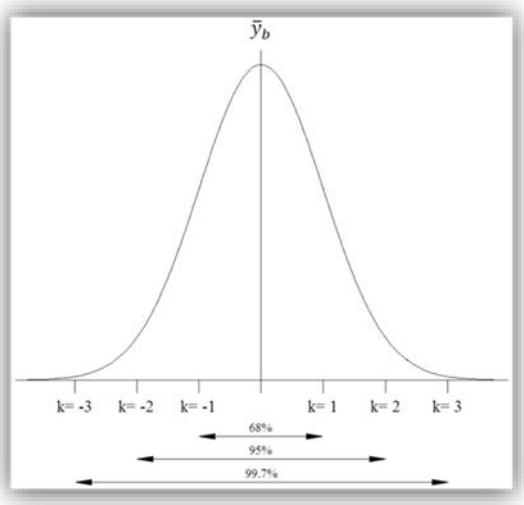
iv. CHARACTERIZING ACCURACY: "Accuracy…includes the concepts of both bias and precision and is *judged with respect to the use to be made of the data*. A measurement process must be unbiased to be capable of producing accurate values…it must be sufficiently precise as well, or else the individual results will be inaccurate due to unacceptable variability."[164]

v. Confidence interval: A range of values symmetric about the bias adjusted mean constructed using a multiple of the standard deviation of the set of measurements and expected to cover the true value with a given level of confidence (likelihood).

$$C.I. = \bar{y}_b \pm k\sigma_y$$

[163] Ehrlich, *Evolution of philosophy and description of measurement* 12 ACCRED. QUAL. ASSUR. 201, 205 (2007).
[164] NIST, *Standard Reference Materials: Handbook for SRM Users*, NISTSP 260-100, 2 (1993).

The likelihood that the interval will overlap the true value is determined by the multiplier of the standard deviation (k), known as a coverage factor, and the underlying distribution. If the underlying distribution is Gaussian (normal) the likelihood associated with $k = 1, 2 \& 3$, is given in the following figure.



One should be very careful with the interpretation of a confidence interval. The focus of the level of confidence is not the true value. That is, the level of confidence does not refer to the probability that the true value lies within the interval. It either does or does not. Rather, the subject of the level of confidence is the sampling procedure. It tells you that based upon the procedure utilized, you will be able to construct an interval that will overlap the true value a given percent of the time. In technical terms, "[t]he confidence reflects the proportion of cases that the confidence interval would contain the true parameter value in a long series of repeated random samples under identical conditions."[165] The confidence interval is based upon frequentist philosophy and the existence of a singular true value.

d) Measurement Interpretation – II: If a measurement value is to be interpretable, we must have a *quantitative* determination of the *systematic* and *random error* associated with the measurement. Absent such information, a measured value is simply a number, the meaning of which we know little about. It has long been understood that no measurement result can be interpreted where only the value of the measurement itself is reported. Proper interpretation of a measured value requires knowledge and incorporation of the measurement's systematic and random error into any reported values.

Unfortunately, as useful as traditional error analysis is, "[i]t is now widely recognized that, when all of the known or suspected components of error have been

[165] ISO, *Statistics — Vocabulary and symbols — Part 1: General statistical terms and terms used in probability*, ISO 3534-1, § 1.28 (2007).

evaluated and the appropriate corrections have been applied, there still remains an uncertainty about the correctness of the stated result, that is, a doubt about how well the result of the measurement represents the value of the quantity being measured."[166] Put simply, it is not possible to know the true value of a measurand or the error of a measurement result and hence how close a measurement result is to the true measurand value.[167]

iii. MEASUREMENT UNCERTAINTY:

a) "[F]or a given measurand and a given result of measurement of it, there is not one value but an infinite number of values dispersed about the result that are consistent with all of the observations and data and one's knowledge of the physical world, and that with varying degrees of credibility can be attributed to the measurand."[168]

b) "The objective of measurement in the Uncertainty Approach is not to determine a true value as closely as possible. Rather, it is assumed that the information from measurement only permits assignment of an interval of reasonable values to the measurand, based on the assumption that no mistakes have been made in performing the measurement. Additional relevant information may reduce the range of the interval of values that can reasonably be attributed to the measurand. However, even the most refined measurement cannot reduce the interval to a single value because of the finite amount of detail in the definition of a measurand. The definitional uncertainty, therefore, sets a minimum limit to any measurement uncertainty. The interval can be represented by one of its values, called a 'measured quantity value.[169]'"

c) Contrary to the traditional approach, then, the measurand is not treated as having a unique "true" value. Instead, the measurand is deemed to consist of a set of "true" values. Measurement uncertainty is a quantitative statement characterizing the dispersion of values that can actually and reasonably be attributed "to a measurand based on the information available including systematic and random effects…and any other factors that may impact the measurement or test process or result."[170] Measurement uncertainty is based upon the Bayesian notion of probability as a measure of degree of belief.

iv. DETERMINING UNCERTAINTY:

a) "When estimating the uncertainty of measurement, all uncertainty components which are of importance in the given situation shall be taken into account using

[166] BIPM, *Evaluation of measurement data — Guide to the expression of uncertainty in measurement (GUM)*, § 0.2, 3.2.2 – 3.2.3 (2008).
[167] BIPM, *Evaluation of measurement data — Guide to the expression of uncertainty in measurement (GUM)*, §3.2.1, 3.3.2 (2008); Ehrlich, *Evolution of philosophy and description of measurement* 12 ACCRED. QUAL. ASSUR. 201, 210 (2007).
[168] BIPM, *Evaluation of measurement data — Guide to the expression of uncertainty in measurement (GUM)*, § 5.2 (2008).
[169] JCGM, *International Vocabulary of Metrology — Basic and General Concepts and Associated Terms (VIM)*, § 0.1 (2008).
[170] JCGM, *International Vocabulary of Metrology — Basic and General Concepts and Associated Terms (VIM)*, § 2.26 (2008); ASTM, *Standard Terminology Relating to Quality and Statistics*, E 456 § 3 (2008).

appropriate methods of analysis."[171]

b) All known systematic effects should be compensated for through application of a correction factor.[172] "It is assumed that a correction (or correction factor) is applied to compensate for each recognized systematic effect that significantly influences the measurement result."[173] Assuming we have determined a systematic error (bias) of *b*, our best estimate of Y would then be:[174]

$$\begin{aligned} Y &= \bar{y} - b + \varepsilon \\ &= y + \varepsilon \end{aligned}$$

c) "Uncertainty of measurement comprises, in general, many components. Some of these components may be evaluated from the statistical distribution of the results of series of measurements and can be characterized by experimental standard deviations. The other components, which can also be characterized by standard deviations, are evaluated from assumed probability distributions based on experience or other information."[175]

d) STANDARD UNCERTAINTY: The total uncertainty associated with any measurement result is typically the result of the combination of several smaller uncertainties associated with particular aspects of the measurement process. Each component of uncertainty that contributes to the uncertainty of a measurement result is known as a standard uncertainty. Each standard uncertainty is expressed and treated as, and may in fact be, a standard deviation.

$$\mu \equiv \sigma$$

e) Relative Standard Uncertainty: The standard uncertainty expressed as a proportion relative to the mean of a set of measurements. It can be useful when combining standard uncertainties or comparing the uncertainty of separate measurements.

$$\mu_r = \frac{\mu_y}{|\bar{y}_b|}$$

f) TYPES OF UNCERTAINTY

[171] ISO, *General requirements for the competence of testing and calibration laboratories*, ISO 17025 § 5.4.6.3 (2005).
[172] JCGM, *Evaluation of measurement data — Guide to the expression of uncertainty in measurement (GUM)*, § 6.3.2 (2008); JCGM, *International Vocabulary of Metrology — Basic and General Concepts and Associated Terms (VIM)*, § 2.53 (2008).
[173] NIST, *Guidelines for Evaluating and Expressing the Uncertainty of NIST Measurement Results*, NIST TN 1297, §5.2, App. D 1.1.6 – 8 (1994).
[174] Eleftheriou, *Measuring performance in analytical measurements* 14 ACCRED. QUAL. ASSUR. 67, 67 (2009).
[175] NIST, Handbook 150 § 1.5.31 Note 2 (2006).

i. TYPE A UNCERTAINTY: Component of measurement uncertainty determined by a statistical analysis of a series of measured quantity values obtained under defined measurement conditions.[176]

1) "A Type A evaluation of standard uncertainty may be based on any valid statistical method for treating data."[177]

a) Standard deviation of the mean of a series of independent observations;

b) Using the method of least squares to fit a curve to data in order to estimate the parameters of the curve and their standard deviations;

c) Carrying out an analysis of variance (ANOVA) in order to identify and quantify random effects in certain kinds of measurements.

ii. TYPE B UNCERTAINTY: Component of measurement uncertainty determined by a method other than the statistical analysis of series of observations.[178] Determination assumes *a priori* distributions based on relevant information and scientific judgment. Examples include information provided by instrument manufacturer, metrological certifications and reference publications.

1) "A Type B evaluation of standard uncertainty is usually based on scientific judgment using all the relevant information available."[179]

a) Previous measurement data;

b) Experience with, or general knowledge of, the behavior and property of relevant materials and instruments;

c) Manufacturer's specifications;

d) Data provided in calibration and other reports;

e) Uncertainties assigned to reference data taken from handbooks;

f) Information associated with the quantity value of a certified reference material;

g) Information about instrumental drift.

[176] JCGM, *International Vocabulary of Metrology — Basic and General Concepts and Associated Terms (VIM)*, § 2.28 (2008); JCGM, *Evaluation of measurement data — Guide to the expression of uncertainty in measurement (GUM)*, § 2.3.2 (2008); NIST, *Guidelines for Evaluating and Expressing the Uncertainty of NIST Measurement Results*, NIST TN 1297, § 2.5 (1994).
[177] NIST, *Guidelines for Evaluating and Expressing the Uncertainty of NIST Measurement Results*, NIST TN 1297, § 3 (1994).
[178] JCGM, *Evaluation of measurement data — Guide to the expression of uncertainty in measurement (GUM)*, § 2.3.3 (2008); JCGM, *International Vocabulary of Metrology — Basic and General Concepts and Associated Terms (VIM)*, § 2.29 (2008).
[179] NIST, *Guidelines for Evaluating and Expressing the Uncertainty of NIST Measurement Results*, NIST TN 1297, § 4.1 (1994); JCGM, *International Vocabulary of Metrology — Basic and General Concepts and Associated Terms (VIM)*, § 2.29 (2008).

iii. "The purpose of the Type A and Type B classification is to indicate the two different ways of evaluating uncertainty components…the uncertainty components resulting from either type are quantified by variances or standard deviations."[180]

1) "Type A evaluations of standard uncertainty components are founded on frequency distributions while Type B evaluations are founded on *a priori* distributions. It must be recognized that in both cases the distributions are models that are used to represent the state of our knowledge."[181]

2) "[T]he GUM approach, and in fact the uncertainty approach in general, are consequences of the Bayesian theory of describing one's state of knowledge about a measurand."[182]

a) "The frequentist theory of inference can be useful for determining certain Type A components of measurement uncertainty, but is not capable of treating most Type B components."[183]

b) "An example of the difficulty of the frequentist theory of inference within the GUM approach is that the frequentist theory is not able to be used to assess the uncertainty of a single measured value when using a measuring instrument, such as a voltmeter. The reason is that the uncertainty here derives from 'nonstatistical' information obtained from the instrument's calibration certificate."[184]

g) "Sources contributing to the uncertainty include, but are not necessarily limited to, the reference standards and reference materials used, methods and equipment used, environmental conditions, properties and condition of the item being tested or calibrated, and the operator."[185]

h) UNCERTAINTY BUDGET: Statement of a measurement uncertainty, of the components of that measurement uncertainty, and of their calculation and combination.[186]

[180] JCGM, *Evaluation of measurement data — Guide to the expression of uncertainty in measurement (GUM)*, § 3.3.4 (2008).
[181] JCGM, *Evaluation of measurement data — Guide to the expression of uncertainty in measurement (GUM)*, § 4.1.6 (2008).
[182] Ehrlich, *Evolution of philosophy and description of measurement* 12 ACCRED. QUAL. ASSUR. 201, 212-213 (2007).
[183] Ehrlich, *Evolution of philosophy and description of measurement* 12 ACCRED. QUAL. ASSUR. 201, 212-213 (2007); Brüchle, *Confidence intervals for experiments with background and small numbers of events* 91 RADIOCHIM. ACTA 71, 71 (2003).
[184] Ehrlich, *Evolution of philosophy and description of measurement* 12 ACCRED. QUAL. ASSUR. 201, 213 (2007).
[185] ISO, *General requirements for the competence of testing and calibration laboratories*, ISO 17025 § 5.4.6.3 Note 1 (2005); JCGM, *Evaluation of measurement data — Guide to the expression of uncertainty in measurement (GUM)*, § 3.3.2 (2008).
[186] JCGM, *International Vocabulary of Metrology — Basic and General Concepts and Associated Terms (VIM)*, § 2.33 (2008).

Uncertainty Source	Type A	Type B	
Calibration			
Ref. Mat.		.052	
Precision	. 080		
Bias	.068		
Combined Uncertainty by Type	.105	.052	
Combined Uncertainty Calibration			.117
Instrumental			
Mechanical Effects	.064		
Electronic Stability	.055		
Detector		.041	
Combined Uncertainty by Type	.084	.041	
Combined Uncertainty Instrumental			.093
Measurement			
Environmental Factors	.101		
Sampling	.112		
Operator	.064		
Measurand Effects		.055	
Combined Uncertainty by Type	.164	.055	
Combined Uncertainty Measurement			.173
Total Uncertainty			
Combined Uncertainty			.229
Expanded Uncertainty (k=2)			± .458

i) COMBINED UNCERTAINTY: The combination of all the standard uncertainties associated with a measurement. The individual standard uncertainties are combined in the same manner as standard deviations. Assuming the standard uncertainties are random and independent, the combined uncertainty is the root sum square of the standard uncertainties. The combined uncertainty is expressed and treated as, and may in fact be, a standard deviation.[187] When determining the combined uncertainty of a measurement it is critical to include all significant components of uncertainty. Failure to do so will cause an underestimate of the uncertainty misleading others to believe that the result is more precise than it actually is. The combined uncertainty can be represented symbolically as:

$$\mu_c = \sqrt{\sum_{i=1}^{n} \mu_i^2}$$

j) EXPANDED UNCERTAINTY: Value obtained when the combined uncertainty is multiplied by a "coverage factor." The expanded uncertainty can be represented symbolically as:[188]

$$U = k\mu_c$$

[187] JCGM, *International Vocabulary of Metrology — Basic and General Concepts and Associated Terms (VIM)*, § 2.31 (2008); JCGM, *Evaluation of measurement data — Guide to the expression of uncertainty in measurement (GUM)*, § 2.3.4 (2008); NIST, *Guidelines for Evaluating and Expressing the Uncertainty of NIST Measurement Results*, NIST TN 1297 § 5 (1994).

[188] JCGM, *Evaluation of measurement data — Guide to the expression of uncertainty in measurement (GUM)*, § 6.2.1 (2008); JCGM, *International Vocabulary of Metrology — Basic and General Concepts and Associated Terms (VIM)*, § 2.35 (2008); NIST, *Guidelines for Evaluating and Expressing the Uncertainty of NIST Measurement Results*, NIST TN 1297 § 6 (1994).

i. A coverage factor is chosen such that when the expanded uncertainty is expressed as part of a complete measurement result it conveys a range of values that can actually and reasonably be attributed to a measurand with a given level of confidence. The level of confidence associated with a given coverage factor is determined by the measurement's underlying distribution. If the underlying distribution is Gaussian (normal) the level of confidence associated with $k = 1.64, 1.96$ & 2.576, is given in the following table.[189]

k	level of confidence
1.64	90%
1.96	95%
2.576	99%

ii. LEVEL OF CONFIDENCE: Probability that the set of true quantity values of a measurand is contained within a specified coverage interval.[190]

1) The level of confidence attributed to a coverage interval is dependent on assumptions regarding the probability distribution associated with a measurement result and its combined standard uncertainty. It is only valid to the extent to which the assumptions may be justified.[191]

2) For a given set of assumptions, the level of confidence provided by an interval is determined by the coverage factor chosen.[192]

k) THE COVERAGE INTERVAL:

i. An "interval containing the set of true quantity values of a measurand with a stated probability, based on the information available."[193] Ordinarily the coverage interval is derived from the expanded uncertainty and is symmetric about the mean so that it can be expressed as:

$$C = \bar{y}_b \pm U \ (99\%)$$

Unlike the confidence interval, the coverage interval is based upon Bayesian philosophy so that it refers directly to the quantity of interest, the "true" value of the measurand. In this context, the level of confidence is the probability, understood as a degree of belief, "that the set of true quantity values of a

[189] JCGM, *International Vocabulary of Metrology — Basic and General Concepts and Associated Terms (VIM)*, § 2.35 - § 2.38 (2008); JCGM, *Evaluation of measurement data — Guide to the expression of uncertainty in measurement (GUM)*, § 2.3.5 - § 2.3.6 (2008).
[190] JCGM, *International Vocabulary of Metrology — Basic and General Concepts and Associated Terms (VIM)*, § 2.37 (2008); JCGM, *Evaluation of measurement data — Guide to the expression of uncertainty in measurement (GUM)*, § 2.3.5 (2008).
[191] JCGM, *Evaluation of measurement data — Guide to the expression of uncertainty in measurement (GUM)*, § 2.3.5 (2008).
[192] JCGM, *Evaluation of measurement data — Guide to the expression of uncertainty in measurement (GUM)*, § 6.3.1 (2008).
[193] JCGM, *International Vocabulary of Metrology — Basic and General Concepts and Associated Terms (VIM)*, § 2.36 (2008).

measurand is contained within the range $\bar{y}_b - U$ to $\bar{y}_b + U$."[194] It should also be noted that the coverage interval need not be symmetric about the mean.[195]

l) SAFETY MARGIN:

i. Another way to account for uncertainty is to subtract "a 'safety margin' from the result to ensure that…the result does not exceed a limit value only because of random effects of the measurement."[196]

ii. The magnitude of the safety margin "depends both on the acceptable risk of committing a type 1 error [false positive] and on the uncertainty of the result."[197]

iii. This is similar to utilization of a one sided confidence interval.[198] Assuming a safety margin *s*, our estimate of Y becomes:

$$Y \geq y - s$$

iv. This would be interpreted to mean that a very small fraction of the distribution of values that could reasonably be attributed to Y would be encompassed by the region $Y < y - s$.[199]

[194] JCGM, *International Vocabulary of Metrology — Basic and General Concepts and Associated Terms (VIM)*, § 2.37, 6.2.1 (2008).
[195] JCGM, *International Vocabulary of Metrology — Basic and General Concepts and Associated Terms (VIM)*, § 2.36 (2008); UKAS, *The Expression of Uncertainty and Confidence in Measurement*, M3003 § 6.7 (2007); NIST, *Guidelines for Evaluating and Expressing the Uncertainty of NIST Measurement Results*, NIST TN 1297 § 6.1 (1994); Eleftheriou, *Measuring performance in analytical measurements* 14 ACCRED. QUAL. ASSUR. 67, 67 (2009); Richter, *Reporting measurement uncertainty in chemical analysis*, 13 ACCRED. QUAL. ASSUR. 113, 113 (2008); Brüchle, *Confidence intervals for experiments with background and small numbers of events* 91 RADIOCHIM. ACTA 71, 71 (2003)..
[196] Kristiansen, *An Uncertainty Budget for the Measurement of Ethanol in Blood by Headspace Gas Chromatography*, 28(6) J. ANAL. TOX. 456 (2004).
[197] Kristiansen, *An Uncertainty Budget for the Measurement of Ethanol in Blood by Headspace Gas Chromatography*, 28(6) J. ANAL. TOX. 456 (2004).
[198] ISO, *Statistics — Vocabulary and symbols — Part I: General statistical terms and terms used in probability*, ISO 3534-1 § 1.29 (2006).
[199] JCGM, *Evaluation of measurement data — Guide to the expression of uncertainty in measurement (GUM)*, § C.2.28 (2008); Garaj, *One-sided Tolerance Factors of Normal Distributions with unknown mean and variability*, 6(2) MEAS. SCI. REV. 12, 14-15 (2006); Kristiansen, *An Uncertainty Budget for the Measurement of Ethanol in Blood by Headspace Gas Chromatography*, 28(6) J. ANAL. TOX. 456 (2004).

d. UNRELIABILITY OF OBSERVATIONS (QUALITATIVE METHODS):

i. Traditional probabilistic measures of unreliability include:[200]

a) FALSE NEGATIVE (TYPE I ERROR) RATE: Percent rejection of true condition.
$FNR = [N_{FN} /(N_{TP} + N_{FN})]$

b) FALSE POSITIVE (TYPE II ERROR) RATE: Percent failure to reject false condition.
$FPR = [N_{FP} /(N_{FP} + N_{TN})]$

c) SENSITIVITY: Percent confirming a true condition.
$S_e = [N_{TP} /(N_{TP} + N_{FN})]$

d) SPECIFICITY: Percent rejecting a false condition.
$S_p = [N_{TN} /(N_{FP} + N_{TN})]$

e) POSITIVE PREDICTIVE VALUE: Percent indicating condition true that are correct.
$P_{pv} = [N_{TP} /(N_{FP} + N_{TP})]$

f) NEGATIVE PREDICTIVE VALUE: Percent indicating condition false that are correct.
$N_{pv} = [N_{TN} /(N_{FN} + N_{TN})]$

	Test Result A	Test Result ¬A	
Condition A	True Positive N_{TP}	False Negative (Type I error) N_{FN}	$N_{TP} + N_{FN}$
Condition ¬A	False Positive (Type II error) N_{FP}	True Negative N_{TN}	$N_{FP} + N_{TN}$
	$N_{TP} + N_{FP}$	$N_{FN} + N_{TN}$	N

g) "The existence of several types of potential error rates makes it absolutely critical for all involved in the analysis to be explicit and precise in the particular rate or rates referenced in a specific setting."[201]

[200] Rios, *Quality assurance of qualitative analysis in the framework of the European project 'MEQUALAN'*, 8 ACCRED. QUAL. ASSUR. 68, 71 (2003); Mil'man, *Uncertainty of Qualitative Chemical Analysis: General Methodology and Binary Test Methods*, 59(12) J. ANAL. CHEM. 1128, 1130-1134, 1136 (2004); Ellison, *Characterizing the performance of qualitative analytical methods: Statistics and terminology*, 24(6) TRENDS ANAL. CHEM. 468, 469-70 (2005); Lewis, *Reliability and Validity: Meaning and Measurement*, 10-11, Presentation to Annual Meeting of the Society for Academic Emergency Medicine (1999); ISO, *Statistics — Vocabulary and symbols — Part I: General statistical terms and terms used in probability*, ISO 3534-1 §§ 1.46, 1.47 (2006); *Handbook of Parametric and Nonparametric Statistical Procedures* 335 (CRC 2007).
[201] NAS, *Strengthening Forensic Science in the United States: A Path Forward*, 4-9 (2009).

h) "It is important for laboratories to check at least the most critical false response rate for a qualitative test."[202]

ii. An alternative approach is to apply Bayes Theorem.[203]

a) BAYES THEOREM: States that the probability of a hypothesis being true given some result is proportional to the probability of the hypothesis being true prior to obtaining the result multiplied by the probability of obtaining the result assuming the hypothesis is true. "Bayesian inference provides a rigorous means of incorporating prior information into a measurement."[204] It can be written as:[205]

$$p(H \mid I) \propto p(I|H)p(H)$$

where

$p(H \mid I)$ = Posterior probability: Probability of H given result I.
$p(H)$ = Prior probability: Independent probability of H prior to result I.
$p(I \mid H)$ = Probability of result I if H true.

b) LIKELIHOOD RATIO: [206] $L(I \mid H) = \frac{p(I \mid H)}{p(I \mid \neg H)}$. This is a measure of the impact of the test result on the likelihood of H, that is of how much the test result has increased or decreased the pretest likelihood of H.

c) POSTERIOR PROBABILITY:[207] $p(H \mid I)$ = Probability (degree of belief) that H is true given test result I.

6. REPORTING RESULTS:

[202] Ellison, *Uncertainties in qualitative testing and analysis*, 5 ACCRED. QUAL. ASSUR. 346, 348 (2000).
[203] Ellison, *Uncertainties in qualitative testing and analysis*, 5 ACCRED. QUAL. ASSUR. 346, 346 (2000); Mil'man, *Uncertainty of Qualitative Chemical Analysis: General Methodology and Binary Test Methods*, 59(12) J. ANAL. CHEM. 1128, 1137-1138 (2004).
[204] Phillips, *Calculation of Measurement Uncertainty Using Prior Information* 103 J. RES. NATL. INST. STAND. TECHNOL. 625, 626 (1998); Ellison, *Quantifying uncertainty in qualitative analysis* 123 ANALYST 1155, 1156 (1998).
[205] Pearl, *Causality: Models Reasoning and Inference* 5 (Cambridge 2001); Estler, *Measurement as Inference: Fundamental Ideas*, 48(2) Annals of the CIRP 611, 618 (1999); Bolstad, *Introduction to Baysian Statistics* 63, 73 (Wiley 2007); Howson, *Scientific Reasoning The Bayesian Approach* 20-21 (Open Court 2006); Leonard, *Bayesian Methods An Analysis for Statisticians and Interdisciplinary Researchers* 76 (Cambridge 1999); Mendenhall, *Mathematical Statistics with Applications*, 64 (PWS-Kent 1990); Brüchle, *Confidence intervals for experiments with background and small numbers of events* 91 RADIOCHIM. ACTA 71, 74-75 (2003).
[206] Ellison, *Quantifying uncertainty in qualitative analysis* 123 ANALYST 1155, 1157-1158 (1998); Pearl, *Causality: Models Reasoning and Inference* 7 (Cambridge 2001); Bolstad, *Introduction to Baysian Statistics* 63, 70 (Wiley 2007); Howson, *Scientific Reasoning The Bayesian Approach* 20-21 (Open Court 2006); Leonard, *Bayesian Methods An Analysis for Statisticians and Interdisciplinary Researchers* 112 (Cambridge 1999).
[207] Ellison, *Characterizing the performance of qualitative analytical methods: Statistics and terminology*, 24(6) TRENDS ANAL. CHEM. 468, 70 (2005); Mil'man, *Uncertainty of Qualitative Chemical Analysis: General Methodology and Binary Test Methods*, 59(12) J. ANAL. CHEM. 1128, 1137-1138 (2004); Ellison, *Quantifying uncertainty in qualitative analysis* 123 ANALYST 1155, 1157-1158 (1998).

a. "Calculations and data transfers shall be subject to appropriate checks in a systematic manner."[208]

   i. "When the experimenter is clearly aware that a gross deviation from prescribed experimental procedure has taken place, the resultant observation should be discarded, whether or not it agrees with the rest of the data."[209]

   ii. OUTLIER: "[A]n observation that appears to deviate markedly in value from other members of the sample in which it appears."[210]

      a) "An outlying observation may be merely an extreme manifestation of the random variability inherent in the data. If this is true, the value should be retained and processed in the same manner as the other observations in the sample."[211]

      b) "On the other hand, an outlying observation may be the result of gross deviation from prescribed experimental procedure or an error in calculating or recording the numerical value", malfunctions or contamination.[212]

      c) "A single result or an entire set of results is suspected to be a statistically invalid result if its deviation either in accuracy or precision from others in the set or other sets, respectively, is greater than can be justified by statistical fluctuations pertinent to a given frequency distribution."[213]

         i. "Outliers should not be excluded on purely statistical evidence until they have been thoroughly investigated and, where possible, the reasons for the discrepancies identified."[214]

         ii. Chauvenet's Criterion (also known as Grubb's test) is a common test for outliers:[215]

$$C < \frac{|\bar{y} - y_{o1}|}{\sigma}$$

[208] ISO, *General requirements for the competence of testing and calibration laboratories*, ISO 17025 § 5.4.7.1 (2005); NIST, Handbook 150 § 5.4.7.1 (2006).
[209] ASTM, *Standard Practice for Dealing With Outlying Observations*, E 178 § 4.1 (2008).
[210] ASTM, *Standard Terminology Relating to Quality and Statistics*, E456 §3 (2008); ISO, *Accuracy (trueness and precision) of measurement methods and results - Part 1: General principles and definitions* § 3.21 (1994).
[211] ASTM, *Standard Practice for Dealing With Outlying Observations*, E 178 § 1.1.1 (2008); ISO, *Reference Materials – General and Statistical Principles for Certification*, ISO Guide 35 § 10.5.5 (2006).
[212] ASTM, *Standard Practice for Dealing With Outlying Observations*, E 178 § 1.1.2 (2008); NIST, *Standard Reference Materials: Handbook for SRM Users*, NIST SP260-100, 79 (1993).
[213] ISO, *Reference Materials – General and Statistical Principles for Certification*, ISO Guide 35 § 10.5.5 (2006).
[214] ISO, *General Requirements for the Competence of Reference Material Producers*, ISO Guide 34 § 5.15.1 (2000); ASTM, *Standard Practice for Dealing With Outlying Observations*, E 178 § 4.3 (2008); NIST, *Standard Reference Materials: Handbook for SRM Users*, NIST SP260-100, 79 (1993); Taylor, *An Intorduction to Error Analysis: The Study of Uncertainties in Physical Measurements*, 166-9 (2nd 1997); Meyer, *Data Analysis: For Scientists and Engineers*, 17 (1975).
[215] ASTM, *Standard Practice for Dealing With Outlying Observations*, E 178 § 6.1 (2008); NIST, *Standard Reference Materials: Handbook for SRM Users*, NIST SP260-100, 80-81 (1993); Taylor, *An Intorduction to Error Analysis: The Study of Uncertainties in Physical Measurements*, 170 (2nd 1997); Meyer, *Data Analysis: For Scientists and Engineers*, 17-18 (1975).

1) The value chosen for C determines the level of confidence of the outlier test.[216]

d) "For qualitative methods, statistical outliers are represented by abnormally high or low frequencies of incorrect responses."[217]

b. RESULT = MEASUREMENT + UNCERTAINTY:

i. Measurement Result: "In general, the result of a measurement is only an approximation or estimate of the value of the specific quantity subject to measurement, that is, the measurand, and thus the result is complete only when accompanied by a quantitative statement of its uncertainty."[218] Moreover, "[i]t is assumed that the result of a measurement has been corrected for all recognized significant systematic effects and that every effort has been made to identify such effects."[219] Accordingly, a complete measurement result consists of the best estimate of the true value of the measurand, typically the bias adjusted mean, accompanied by the expanded uncertainty and its associated level of confidence.

$$Y = \bar{y}_b \pm U\ (99\%)$$

This is interpreted to mean that the best estimate of the value attributable to the measurand Y is $\bar{y}_b$, and that $\bar{y}_b - U$ to $\bar{y}_b + U$ is the range of values that could actually be attributed to Y with a 99% level of confidence. Note that the coverage interval is identical to the measurement result.

ii. "The result of a measurement cannot be correctly evaluated without knowing its uncertainty."[220]

a) "A quantitative analysis is not a great deal of use unless there is some estimation of how prone to error the analytical procedure is. Simply accepting the analytical result could lead to rejection or acceptance…on the basis of a faulty analysis."[221]

iii. "When reporting the result of a measurement of a physical quantity, it is obligatory that some quantitative indication of the quality of the result be given so that those who use it can assess its reliability. Without such an indication, measurement results cannot

[216] ASTM, *Standard Practice for Dealing With Outlying Observations*, E 178 § 6 (2008); Taylor, *An Intorduction to Error Analysis: The Study of Uncertainties in Physical Measurements*, 166-170, App. A (2nd 1997).
[217] Ellison, *Characterizing the performance of qualitative analytical methods: Statistics and terminology*, 24(6) TRENDS ANAL. CHEM. 468, 475 (2005).
[218] NIST *Guidelines for Evaluating and Expressing the Uncertainty of NIST Measurement Results*, NIST TN 1297 § 2.1 (1994); BIPM, *Evaluation of measurement data — Guide to the expression of uncertainty in measurement (GUM)*, § 7.1.4 (2008).
[219] BIPM, *Evaluation of measurement data — Guide to the expression of uncertainty in measurement (GUM)*, § 3.2.4 (2008); NIST, *Guidelines for Evaluating and Expressing the Uncertainty of NIST Measurement Results*, NIST TN 1297 § 5.2 (1994).
[220] Desimoni, *About considering both false negative and false-positive errors when assessing compliance and non-compliance with reference values given in compositional specifications and statutory limits*, 13 ACCRED. QUAL. ASSUR. 653, 653 (2008).
[221] Watson, *Pharmaceutical Analysis - A Textbook for Pharmacy Students and Pharmaceutical Chemists*, 2 (2nd ed. Elsevier 2005)

be compared, either among themselves or with reference values given in a specification or standard."[222]

a) "Requirements for measurement accuracy translate into a need to know not only the results of measurements but the uncertainties associated with the results."[223]

b) "In general, the result of a measurement is only an approximation or estimate of the value of the measurand and thus is complete only when accompanied by a statement of the uncertainty of that estimate."[224]

c) "Measurement uncertainty is an integral part of a measurement result. Without a statement of uncertainty a measurement result is not complete. Concluding about compatibility with other measurement results obtained for the same measurand or with compliance limits is not possible and the measurement result does therefore, not serve its purpose."[225]

iv. Reports of result must include:[226]

a) Test method – Description of how test was made;

b) Calibration results – When an instrument has been repaired or adjusted the calibration results before and after repair or adjustment are reported;

c) Standards used – Identification of and traceability to national standards;

d) Quantitative methods – Description of calculations of measurement result and its uncertainty from the experimental observations and input data;

i. Include all corrections and constants used in the analysis and their sources;

e) Estimated measurement uncertainty –

i. List all uncertainty components and document fully how they were evaluated;

ii. Coverage factor and estimated confidence interval.

[222] JCGM, *Evaluation of measurement data — Guide to the expression of uncertainty in measurement (GUM)*, § 0.1 (2008).
[223] Ehrlich, *Metrological Timelines in Traceability*, 103 J. Res. Natl. Inst. Stand. Technol. 93, 94 (1998).
[224] JCGM, *Evaluation of measurement data — Guide to the expression of uncertainty in measurement (GUM)*, § 3.1.2 (2008); NIST, *Guidelines for Evaluating and Expressing the Uncertainty of NIST Measurement Results*, NIST TN 1297 § 2.1 (1994); Brüchle, *Confidence intervals for experiments with background and small numbers of events* 91 RADIOCHIM. ACTA 71, 71 (2003).
[225] Richter, *Reporting measurement uncertainty in chemical analysis*, 13 ACCRED. QUAL. ASSUR. 113, 113 (2008).
[226] ISO, *General requirements for the competence of testing and calibration laboratories*, ISO 17025 § 5.10.3 (2005); NIST, *Recommended Standard Operations Procedures for Preparation of Test/Calibration Reports*, SOP-1, § 2 (2003); JCGM, *Evaluation of measurement data — Guide to the expression of uncertainty in measurement (GUM)*, § 7.1.4 (2008); NAS, *Strengthening Forensic Science in the United States: A Path Forward*, S-15, 6-3 (2009).

v. MEASUREMENT + UNCERTAINTY:

a) "It is assumed that the result of a measurement has been corrected for all recognized significant systematic effects and that every effort has been made to identify such effects."[227]

b) COVERAGE INTERVAL APPROACH:

i. "State the result of the measurement as $Y = y \pm U$ and give the units of y and U."[228]

ii. Give the value of λ used to obtain U ($U = \lambda\mu_c$).[229]

iii. "Give the approximate level of confidence associated with the interval $y \pm U$ and state how it was determined."[230]

iv. The expanded uncertainty is generally reported with approximately a 95% - 99% level of confidence.[231]

c) SAFETY MARGIN APPROACH:

i. State the result of the measurement as $Y \geq y - s$ and give the units of y and s.

ii. Give the approximate level of confidence associated with the region $Y \geq y - s$ and state how it was determined.

vi. Measurement Interpretation – III: For even the most carefully performed measurement, a unique "true" value for a measurand can never be determined. All that can ever be given is a set of values, all of which may actually and reasonably be assigned as "true" values. If a measurement value is to be interpretable, it must be *corrected for bias* and accompanied by a *quantitative estimate of its uncertainty*. Absent such information, a measured value is simply a number, the meaning of which we know little about.

"Knowledge of the uncertainty associated with measurement results is essential to the interpretation of the results. Without quantitative assessments of uncertainty, it is impossible to decide whether observed differences between results reflect more than experimental variability, whether test items comply with specifications, or whether laws based on limits have been broken. Without information on uncertainty, there is a

[227] JCGM, *Evaluation of measurement data — Guide to the expression of uncertainty in measurement (GUM)*, § 3.2.4 (2008); NIST, *Guidelines for Evaluating and Expressing the Uncertainty of NIST Measurement Results*, NIST TN 1297 §5.2, App. D 1.1.6 – 8 (1994).
[228] JCGM, *Evaluation of measurement data — Guide to the expression of uncertainty in measurement (GUM)*, § 7.2.3 (2008).
[229] JCGM, *Evaluation of measurement data — Guide to the expression of uncertainty in measurement (GUM)*, § 7.2.3 (2008).
[230] JCGM, *Evaluation of measurement data — Guide to the expression of uncertainty in measurement (GUM)*, § 7.2.3 (2008).
[231] JCGM, *Evaluation of measurement data — Guide to the expression of uncertainty in measurement (GUM)*, § 3.3.7, 6.3 (2008); NIST, *Good Laboratory Practice for Rounding Expanded Uncertainties and Calibration Values*, GLP-9, 1 (2003); NIST, *Guidelines for Evaluating and Expressing the Uncertainty of NIST Measurement Results*, NIST TN 1297 §6.2 – 6.3 (1994); Richter, *Reporting measurement uncertainty in chemical analysis*, 13 ACCRED. QUAL. ASSUR. 113, 113 (2008); UKAS, *The Expression of Uncertainty and Confidence in Measurement*, M3003 § 6.1 – 6.4 (2007).

risk of misinterpretation of results. Incorrect decisions taken on such a basis may result in unnecessary expenditure in industry, incorrect prosecution in law, or adverse health or social consequences."[232]

c. RESULTS OF QUALITATIVE TEST OBSERVATIONS:

i. "The most common, and probably the most useful, form of data treatment in method-validation studies for qualitative tests is the calculation and reporting of either specificity and sensitivity or false positive and negative error rates."[233]

ii. Using Bayes theorem, "[t]he scientist can testify to the value of their evidence by quoting a likelihood ratio value obtained from a particular procedure."[234]

F. SCIENTIFIC STANDARDS: "Measurement is one of the basic tools humanity uses to understand the environment and compare quality. International standardization was established and National laboratories were founded in every advanced society to control this basic measurement need."[235]

1. STANDARD: Document, established by consensus and approved by a recognized body, that provides, for common and repeated use, rules, guidelines or characteristics for activities or their results, aimed at the achievement of the optimum degree of order in a given context."[236]

a. "Standards should be based on the consolidated results of science, technology and experience, and aimed at the promotion of optimum community benefits."[237]

b. Types

i. Basic Standard: Standard that has a wide-ranging coverage or contains general provisions for one particular field.[238]

ii. Testing Standard: Standard that is concerned with test methods, sometimes supplemented with other provisions related to testing, such as sampling, use of statistical methods, sequence of tests.[239]

iii. Process Standard: Standard that specifies requirements to be fulfilled by a process, to establish its fitness for purpose.[240]

---

[232] ISO, *Guidance for the use of repeatability, reproducibility and trueness estimates in measurement uncertainty estimation*, ISO/TS 21748 DRAFT REVISION, v (2009); ISO, *Guidance for the use of repeatability, reproducibility and trueness estimates in measurement uncertainty estimation*, ISO/TS 21748, v (2004).
[233] Ellison, *Characterizing the performance of qualitative analytical methods: Statistics and terminology*, 24(6) TRENDS ANAL. CHEM. 468, 470 (2005).
[234] Ramos, *Information-theoretical comparison of likelihood ratio methods of forensic evidence evaluation*, presented at the THIRD INT. SYM. ON INFO. ASSURANCE AND SEC. (2007); Evett, *A Model for Case Assessment and Interpretation* 38(3) SCI. & JUSTICE 151 (1998).
[235] Eleftheriou, *Measuring performance in analytical measurements* 14 ACCRED. QUAL. ASSUR. 67, 67 (2009).
[236] ISO, *Standardization and related activities — General vocabulary*, ISO 2 § 3.2 (2004).
[237] ISO, *Standardization and related activities — General vocabulary*, ISO 2 § 3.2 Note (2004).
[238] ISO, *Standardization and related activities — General vocabulary*, ISO 2 § 5.1 (2004).
[239] ISO, *Standardization and related activities — General vocabulary*, ISO 2 § 5.3 (2004).
[240] ISO, *Standardization and related activities — General vocabulary*, ISO 2 § 5.5 (2004).

iv. Terminology Standard: Define words permitting parties to use a common, clearly understood language.[241]

v. Standard on Data to be Provided: Standard that contains a list of characteristics for which values or other data are to be stated for specifying the product, process or service.[242]

c. ACKNOWLEDGED RULE OF TECHNOLOGY: Technical provision acknowledged by a majority of representative experts as reflecting the state of the art.[243]

i. STATE OF THE ART: Developed stage of technical capability at a given time as regards products, processes and services, based on the relevant consolidated findings of science, technology and experience.[244]

ii. "A normative document on a technical subject, if prepared with the cooperation of concerned interests by consultation and consensus procedures, is presumed to constitute an acknowledged rule of technology at the time of its approval."[245]

iii. "Voluntary consensus standards are heavily peer-reviewed before they even come into existence."[246]

d. UTILITY:

i. "Standards provide the foundation against which performance, reliability, and validity can be assessed. Adherence to standards reduces bias, improves consistency, and enhances the validity and reliability of results. Standards reduce variability resulting from the idiosyncratic tendencies of the individual examiner…They make it possible to replicate and empirically test procedures and help disentangle method errors from practitioner errors."[247]

ii. "Standards ensure desirable characteristics of services and techniques such as quality, reliability, efficiency, and consistency among practitioners."[248]

iii. "[S]tandards are crucial to every form of scientific and industrial process."[249]

---

[241] Breitenberg, Office of Standards Code and Information, NIST, *The ABC's of Standards-Related Activities in the United States*, NBSIR 87-3576 (1987); ISO, *Standardization and related activities — General vocabulary*, ISO 2 § 5.2 (2004).
[242] ISO, *Standardization and related activities — General vocabulary*, ISO 2 § 5.8 (2004).
[243] ISO, *Standardization and related activities — General vocabulary*, ISO 2 § 1.5 (2004).
[244] ISO, *Standardization and related activities — General vocabulary*, ISO 2 § 1.4 (2004).
[245] ISO, *Standardization and related activities — General vocabulary*, ISO 2 § 1.5 Note (2004).
[246] Lentini, *Forensic Science Standards: Where They Come From and How They Are Used*, 1 FOR. SCI. POL. MGMT. 10, 10 (2009).
[247] NAS, *Strengthening Forensic Science in the United States: A Path Forward*, 7-7 (2009).
[248] NAS, *Strengthening Forensic Science in the United States: A Path Forward*, 7-1 (2009).
[249] Breitenberg, Office of Standards Code and Information, NIST, *The ABC's of Standards-Related Activities in the United States*, NBSIR 87-3576 (1987).

e. "Typically standards are enforced through systems of accreditation and certification, wherein independent examiners and auditors test and audit the performance, policies, and procedures of both laboratories and service providers."[250]

2. ISO 17025: GENERAL REQUIREMENTS FOR THE COMPETENCE OF TESTING AND CALIBRATION LABORATORIES.

a. The Gold Standard: "This International Standard specifies the general requirements for the competence to carry out tests and/or calibrations, including sampling. It covers testing and calibration performed using standard methods, non-standard methods, and laboratory-developed methods."[251]

i. Competence: Ability of a laboratory to conduct tests and perform calibrations in accordance with the specified standards and to produce accurate, proper, fit for purpose, technically valid data and test and calibration results.[252]

b. "This International Standard is applicable to all organizations performing tests and/or calibrations…[and] all laboratories regardless of the number of personnel or the extent of the scope of testing and/or calibration activities."[253]

c. "This international standard forms the basis for international laboratory accreditation."[254]

V. FORENSIC METROLOGY

A. FORENSIC METROLOGY: "Forensic Metrology is the application of measurements and hence measurement standards to the solution and prevention of crime."[255]

1. "Legal metrology is an internationally coordinated activity that aims to ensure the reliability of measurements that might be the subject of dispute in law. It aims to standardize the use of measurement units, to provide, or facilitate the provision of traceable measurement standards and to evaluate and approve certain types of measuring equipment."[256]

2. "The need for a reliable world metrology system is driven not only by trade requirements but equally by societal requirements. Improvement of the quality of life is highly served by reliable, traceable and more accurate measurements, particularly in areas such as…forensics and security."[257]

[250] NAS, *Strengthening Forensic Science in the United States: A Path Forward*, 7-1 – 7-2 (2009).
[251] ISO, *General requirements for the competence of testing and calibration laboratories*, ISO 17025 § 1.1 (2005).
[252] NIST, Handbook 150 § 1.5.8 (2006).
[253] ISO, *General requirements for the competence of testing and calibration laboratories*, ISO 17025 § 1.2 (2005); NIST, Handbook 150, *v-vi* (2006).
[254] UKAS, *The Expression of Uncertainty and Confidence in Measurement*, M3003 § 1.1 (2007).
[255] Sharp, *Measurement Standar*ds, in *Measurement, Instrumentation, and Sensors Handbook* §5.2 (1999).
[256] King, *Chemical measurement and the law: metrology and quality issues*, 6 ACCRED. QUAL. ASSUR. 236, 241 (2001).
[257] Kaarls, *Metrology, essential to trade, industry and society*, 12 ACCRED. QUAL. ASSUR. 423, 435 (2007).

3. Forensic metrology is practiced around the world.[258]

B. FORENSIC WEIGHTS AND MEASURES

1. REFERENCE MATERIALS AND STANDARDS

a. "Access to reference materials and collections is essential to crime laboratory efforts to identify and assign values to materials, calibrate instruments [and] assess measurement methods"[259] as well as to assure the validity of qualitative test results.[260]

b. "Appropriate reference material(s) shall be used for qualitative and quantitative procedures. Traceability of the reference material is required."[261]

i. CERTIFIED REFERENCE MATERIAL: A reference material, accompanied by a certificate, one or more of whose property values are certified by a procedure that establishes traceability to an accurate realization of the unit in which the property values are expressed, and for which each certified value is accompanied by an uncertainty at a stated level of confidence.[262]

a) "A certified reference material…suitable for the preparation of a standard to which calibration material can be compared, must be certified by a method generally recognized by the scientific community as one that validates the CRM for this purpose."[263]

ii. REFERENCE STANDARD: A standard, generally having the highest metrological quality available at a given location or in a given organization, from which measurements made there are derived.[264]

c. Adequacy and documentation of references:

i. "Clear documentation of the [reference material] and its property value(s) should be available, preferably as a certificate ([certified reference material])."[265]

ii. "The quality of standard materials and reagents should be adequate for the procedure used. Lot/batch numbers of standard materials and critical reagents should be recorded. All critical reagents should be tested for their reliability. Standard materials and reagents should be labeled with: name; concentration, where appropriate; preparation

[258] Sharp, *Measurement Standar*ds, in *Measurement, Instrumentation, and Sensors Handbook* §5.2 (1999).
[259] NIST, *1999 Survey of Forensic Reference Materials*, NISTIR 6518, 1 (2000).
[260] Gonzalez-Rodriguez, *Emulating DNA: Rigorous Quantification of Evidential Weight in Transparent and Testable Forensic Speaker Recognition* 15(7) IEEE TRANS. AUDIO SPEECH LANGUAGE PROCESSING 2104, 2104 (2007); Reeder, *Impact of DNA Typing on Standards and Practice in the Forensic Community* 123 ARCH. PATH. LAB. MED. 1063 (1999).
[261] SWGDRUG, *Recommendations (Minimum Standards)*, 31 (2008).
[262] ASTM, *Standard Terminology Relating to Forensic Science*, §4 E 1732 (2005); Epstein, *The Use of Certified Reference Materials in Forensic QA*, Presented at 13th INTERPOL Forensic Science Symposium, (2001).
[263] SOFT/AAFS, *Forensic Toxicology Laboratory Guidelines*, § 9.3.1 (2006).
[264] ASCLD/LAB – International, *Traceability Discussion*, 2 (2004).
[265] Epstein, *The Use of Certified Reference Materials in Forensic QA*, Presented at 13th INTERPOL Forensic Science Symposium, (2001).

date and or expiry date; identity of preparer; storage conditions, if relevant; hazard warning, where necessary."[266]

iii. "Reference collections of data or items/materials encountered in casework which are maintained for identification, comparison or interpretation purposes (eg mass spectra, motor vehicle paints or headlamp lenses, drug samples, typewriter printstyles, wood fragments, bullets, cartridges, DNA profiles, frequency databases) should be fully documented, uniquely identified and properly controlled."[267]

d. "[L]aboratories may obtain certified reference material from NIST…or from another national metrology institute."[268]

C. FORENSIC MEASUREMENT AND TESTING PROCESS

1. TEST METHOD: Defined technical procedure to determine one or more specified characteristics of a material or product.[269]

a. "All methods shall be fully documented including procedures for quality control, and, where appropriate, the use of reference materials."[270]

2. OBJECTIVE TEST: "A test which having been documented and validated is under control so that it can be demonstrated that all appropriately trained staff will obtain the same results within defined limits. These defined limits relate to expressions of degrees of probability as well as numerical values."[271]

a. "Visual inspection, qualitative examinations and computer simulations are included in the definition of objective test."[272]

b. "It is anticipated that the majority of the work carried out in forensic testing laboratories will be capable of satisfying the definition of an objective test."[273]

c. "Objective tests will be controlled by: documentation of the test; validation of the test; training and authorization of staff; maintenance of equipment; and where appropriate by; calibration of equipment; use of appropriate reference materials; provision of guidance for interpretation; checking of results; testing of staff proficiency; recording of equipment/test performance."[274]

[266] ILAC, *Guideline for Forensic Science Laboratories*, ILAC G19, § 5.4.2d (2002).
[267] ILAC, *Guideline for Forensic Science Laboratories*, ILAC G19, § 5.6.3.2 (2002).
[268] FQS-I, *Traceability*, FRAP-4, § 2.1 (2008); Vallone, *Development and usage of a NIST standard reference material for real time PCR quantitation of human DNA* FOR. SCI. INT.: GENETICS SUPP. SERIES 1, 80 (2008).
[269] ASTM, *Standard Terminology Relating to Forensic Science*, §4 E 1732 (2005).
[270] FQS-I, *Forensic Requirements for Accreditation*, FRA-1, § 5.4.1 (2008).
[271] ILAC, *Guideline for Forensic Science Laboratories*, ILAC G19, § 3 (2002).
[272] ILAC, *Guideline for Forensic Science Laboratories*, ILAC G19, § 3 (2002).
[273] FQS-I, *Forensic Requirements for Accreditation*, FRA-1, § 3 (2008).
[274] ILAC, *Guideline for Forensic Science Laboratories*, ILAC G19, § 3 (2002).

3. VALIDATION: Confirmation by examination and provision of objective evidence that the particular requirements for a specific intended use are fulfilled.[275]

    a. "To confirm the validity of a method or process for a particular purpose (e.g., for a forensic investigation), validation studies must be performed."[276]

    b. "All technical procedures used by a forensic science laboratory must be fully validated before being used on casework."[277]

        i. "The reliability of analytical findings is a matter of great importance in forensic and clinical toxicology, as the results may have wide legal consequences or lead to the wrong treatment of a patient. So, at the very least, routine analytical methods have to be validated."[278]

        ii. "Establishing fitness-for-purpose is necessary before analytical results can be relied on for important legal decisions… Given the serious penalties associated with conviction, the entire analytical system must be demonstrated fit-for-purpose."[279]

    a. "The contribution of random and systematic errors to method result uncertainty shall be assessed and the expanded uncertainty derived for quantitative methods."[280]

    b. "In validating test methods, the following issues (among others) may need to be determined, as appropriate: matrix effects; interferences; sample homogeneity; concentration ranges; specificity; stability of measured compounds; linearity range; population distribution; precision; measurement uncertainty."[281]

D. FORENSIC QUALITY ASSURANCE: "Forensic quality control results from an appropriate balance between instrumental and protocol considerations. Many jurisdictions, unfortunately, expend significant effort on instrument selection and testing while giving little thought to the analytical protocol. Forensic integrity results from the balanced contribution of all elements affecting measurement results."[282]

    1. TRACEABILITY: Property of the result of a measurement or value of a standard whereby it can be related with a stated uncertainty, to stated references, usually national or international standards (i.e. through an unbroken chain of comparisons).[283]

[275] ASTM, *Standard Terminology Relating to Forensic Science*, §4 E 1732 (2005).
[276] NAS, *Strengthening Forensic Science in the United States: A Path Forward*, 113 (2009).
[277] ILAC, *Guideline for Forensic Science Laboratories*, ILAC G19, § 5.4.5.1 (2002); FQS-I, *Forensic Requirements for Accreditation*, FRA-1, § 5.4.2 (2008).
[278] Westphal, *Development of a validated method for the simultaneous determination of amphetamine, methamphetamine and methylenedioxyamphetamines (MDA, MDMA, MDEA) in serum by GC-MS after derivatisation with perfluorooctanoyl chloride*, 12 ACCRED. QUAL. ASSUR. 335, 340 (2007).
[279] Gullberg, *Estimating the measurement uncertainty in forensic breath-alcohol analysis*, 11 ACCRED. QUAL. ASSUR. 562, 562 (2006).
[280] SWGDRUG, *Recommendations (Minimum Standards)*, 34 (2008).
[281] ILAC, *Guideline for Forensic Science Laboratories*, ILAC G19, § 5.4.5.1 (2002).
[282] Gullberg, *Methodology and Quality Assurance in Forensic Breath Alcohol Analysis*, 12 For. Sci. Rev. 49, 56 (2000).
[283] ASTM, *Standard Terminology Relating to Forensic Science*, §4 E 1732 (2005).

a. Accuracy and Reliability:

   i. "It is not possible to determine a reliable result and its uncertainty if there is no traceability of the measurement to a standard with known uncertainty. So for reliable results, traceability of each…measurement to a national standard…(or the SI) is essential."[284]

   ii. "Especially for legal purposes, traceability is an essential requirement, which however is not always fulfilled."[285]

   iii. "Traceability to authoritative reference standards is an important and often overlooked element in forensic…analysis."[286]

   iv. "[B]ias can be corrected when traceability is established."[287]

b. "It is a fundamental requirement that the results of all…calibrations required to support accredited tests shall be traceable to national and international standards of measurement."[288]

   i. "ISO/IEC 17025 details the specific requirements for traceablilty to be met by testing and calibration laboratories."[289]

   ii. For the purpose of assuring traceability, testing laboratories that perform calibration only for themselves may calibrate its own equipment if the appropriate requirements of NIST Handbook 150 have been met.[290]

c. DOCUMENTATION: "Accounting for and documenting traceability…is an important element of quality control."[291]

   i. "The laboratory or calibration provider must document the measurement process or system used to demonstrate traceability and provide a description of the chain of comparisons/calibrations that were used to establish a connection to a particular stated reference."[292]

   ii. "To support traceability, the laboratory records for each step in the chain shall include: A clear description of the quantity being measured; Specific information pertaining to the equipment subject to traceability; A complete description of the measurement equipment or working standard used to perform the measurement; A complete specification of the stated reference at the time the measurement system or working standard was compared to it; A stated measurement result or value, with reference to

[284] Knopf, *Traceability system for breath-alcohol measurements in Germany*, OIML Bulletin XLVIII(2), 17 (2007).
[285] Bich, *Interdependence between measurement uncertainty and metrological traceability* ACCRED. QUAL. ASSUR. (IN PRESS - 2009).
[286] Gullberg, *Methodology and Quality Assurance in Forensic Breath Alcohol Analysis*, 12 For. Sci. Rev. 49, 59 (2000).
[287] Gullberg, *Estimating the measurement uncertainty in forensic breath-alcohol analysis*, 11 ACCRED. QUAL. ASSUR. 562, 563 (2006).
[288] FQS-I, *Traceability*, FRAP-4 (2008).
[289] FQS-I, *Traceability*, FRAP-4 (2008).
[290] FQS-I, *Traceability*, FRAP-4 § 2.2 (2008).
[291] Gullberg, *Estimating the measurement uncertainty in forensic breath-alcohol analysis*, 11 ACCRED. QUAL. ASSUR. 562, 568 (2006).
[292] ASCLD/LAB – International, *Measurement Traceability Policy*, 1 (2004).

International System of Units (SI) where possible; A documented uncertainty of measurement and a description of the process used to develop it; Appropriate intervals for re-calibration or calibration checks; Information establishing the competence of the calibration laboratory and/or in-house personnel involved."[293]

iii. "The uncertainty of measurement for each step in the traceability chain must be determined and stated."[294]

2. CALIBRATION: The set of operations that establishes, under specified conditions, the relationship between values indicated by a measuring instrument or measuring system or values represented by a material, and the corresponding known values of measurement.[295]

a. When required:

i. "All equipment used for tests and/or calibrations, including equipment for subsidiary measurements (e.g. for environmental conditions) having a significance effect on the accuracy or validity of the result of the test, calibration or sampling shall be calibrated before being put into service."[296]

ii. "Calibration must be performed…at appropriate intervals thereafter."[297]

iii. "It will normally be necessary to check instrument calibration after any shut down, whether deliberate or otherwise, and following service or other substantial maintenance. In general, calibration intervals should not be less stringent than manufacturers' recommendations."[298]

b. Calibration defines the valid range of measurement:

i. "The concentration of the calibrators should be such that they bracket the anticipated concentration of the specimen(s)."[299]

ii. "The range of the calibration curve should cover the range of concentrations expected in the samples. The calibration curve should not normally be extrapolated beyond the lowest or highest standard solutions."[300]

3. QUALITY ASSURANCE PROGRAM

a. "Forensic laboratories should establish routine quality assurance and quality control procedures to ensure the accuracy of forensic analyses and the work of forensic practitioners. Quality control procedures should be designed to identify mistakes, fraud, and bias; confirm

[293] ASCLD/LAB – International, *Measurement Traceability Policy*, 1 (2004).
[294] ASCLD/LAB – International, *Traceability Discussion*, 2 (2004).
[295] ASTM, *Standard Terminology Relating to Forensic Science*, §4 E 1732 (2005).
[296] ASCLD/LAB – International, *Traceability Discussion*, 1 (2004).
[297] ASCLD/LAB – International, *Traceability Discussion*, 2 (2004).
[298] ILAC, *Guideline for Forensic Science Laboratories*, ILAC G19, § 5.6.1 (2002).
[299] SOFT/AAFS, *Forensic Toxicology Laboratory Guidelines*, § 8.3.6 (2006).
[300] Flanagan, *Fundamentals of Analytical Toxicology* 357 (Wiley 2007).

the continued validity and reliability of standard operating procedures and protocols; ensure that best practices are being followed; and correct procedures and protocols that are found to need improvement."[301]

b. ACCREDITATION: "[P]rocedure by which an authoritative body gives formal recognition that a body or person is competent to carry out specific tasks."[302]

i. "Accreditation deals directly with the ability of a laboratory to provide quality forensic science service."[303]

ii. Accreditation of forensic laboratories must be mandatory.[304]

iii. "Accreditation is part of a laboratory's quality assurance program which should also include proficiency testing."[305]

c. PROFICIENCY TESTING:

i. Proficiency testing is an important aspect of ensuring that forensic laboratories can satisfy minimum standards.[306]

ii. "An effective means for a forensic science laboratory to monitor its performance, both against its own requirements and against the performance of peer laboratories, is to take part in proficiency testing programs."[307]

E. FORENSIC MEASUREMENT/OBSERVATION INTERPRETATION

1. MEASUREMENT/OBSERVATION RESULT

a. "As with all other scientific investigations, laboratory analyses conducted by forensic scientists are subject to measurement error."[308]

b. "Although some forensic scientists may find the notion of 'error' unsettling, it is a reality of measurement that must be appreciated…Only when measurement 'error' is acknowledged and properly estimated can…analytical goals [be] achieved."[309]

[301] NAS, *Strengthening Forensic Science in the United States: A Path Forward*, 7-19 (2009).
[302] ASTM, *Standard Terminology Relating to Forensic Science*, E 1732 § 4.1 (2005).
[303] O'Dell, *A quality assurance system for DNA testing* 2(1) FOR. SCI. J. __ (2003).
[304] NAS, *Strengthening Forensic Science in the United States: A Path Forward*, 7-18 (2009); FQS-I, *Position Statement Regarding NAS Report* (2009).
[305] ASCLD/LAB – International, *Lab International Accreditation Program*, 3 (2006).
[306] SWGDRUG, *Recommendations (Minimum Standards)*, 25 (2008); Gullberg, *Results of a Proposed Breath Alcohol Proficiency Test Program*, 51(1) J. For. Sci. 168,168 (2006).
[307] FQS-I, *Proficiency Testing*, FRAP-2, 3 (2009).
[308] NAS, *Strengthening Forensic Science in the United States: A Path Forward*, 4-5 (2009); Jones, *Dealing with Uncertainty in Chemical Measurements*, 14(1) NEWSL. OF THE INT. ASSOC. FOR CHEM. TEST. 6 (2003).
[309] Gullberg, *Estimating the measurement uncertainty in forensic breath-alcohol analysis*, 11 ACCRED. QUAL. ASSUR. 562, 563 (2006).

c. "All analytical results, regardless of context, protocol or instrumentation, possess uncertainty…all measurement results are approximations. This is acceptable…so long as the limits of uncertainty are known and acceptable."[310]

2. FORENSIC DETERMINATION OF UNCERTAINTY AND UNRELIABILITY

a. FUNDAMENTAL TO UNDERSTANDING OF TEST RESULT

i. "[T]he most important questions that any *soi-disant* expert must be asked, and be able satisfactorily to answer, are what is the scientific basis of your claim, and what is your error rate?"[311]

ii. "Only when measurement 'error' is acknowledged and properly estimated can…analytical goals [be] achieved."[312]

iii. "Knowledge of the uncertainty associated with measurement results is essential to the interpretation of the results. Without quantitative assessments of uncertainty, it is impossible to decide…whether laws based on limits have been broken. Without information on uncertainty, there is a risk of misinterpretation of results. Incorrect decisions taken on such a basis may result in…incorrect prosecution in law."[313]

iv. "The assessment of the accuracy of the conclusions from forensic analyses and the estimation of relevant error rates are key components of the mission of forensic science."[314]

a) "Many would consider inadequate statistical thought in experimental design and data analysis to be unethical scientific practice. Modern analytical systems must be shown to have sufficient accuracy, precision [and] uncertainty estimates"[315]

b. ISO/NIST METHODOLOGY: Forensic determination and reporting of uncertainty is governed by the requirements of ISO 17025[316] and NIST 1297.[317]

c. FORENSIC DETERMINATION OF UNCERTAINTY:

i. BEST ESTIMATE OF MEASURAND VALUE:

---

[310] Gullberg, *Methodology and Quality Assurance in Forensic Breath Alcohol Analysis*, 12 For. Sci. Rev. 49, 50 (2000).
[311] Gonzalez-Rodriguez, *Emulating DNA: Rigorous Quantification of Evidential Weight in Transparent and Testable Forensic Speaker Recognition* 15(7) IEEE TRANSACTIONS ON AUDIO, SPEECH, AND LANGUAGE PROCESSING 2104, 2113 (2007).
[312] Gullberg, *Estimating the measurement uncertainty in forensic breath-alcohol analysis*, 11 ACCRED. QUAL. ASSUR. 562, 563 (2006).
[313] ISO, *Guidance for the use of repeatability, reproducibility and trueness estimates in measurement uncertainty estimation*, ISO/TS 21748 DRAFT REVISION, v (2009); ISO, *Guidance for the use of repeatability, reproducibility and trueness estimates in measurement uncertainty estimation*, ISO/TS 21748, v (2004).
[314] NAS, *Strengthening Forensic Science in the United States: A Path Forward*, 4-9 (2009).
[315] Gullberg, *Statistical Applications in Forensic Toxicology*, Medical-Legal Aspects of Alcohol, p. 457, 458 (James Garriott ed., 5th ed. 2009).
[316] ILAC, *Guideline for Forensic Science Laboratories*, ILAC G19, § 5.10 (2002); ASCLD/LAB – *International, Estimating Uncertainty of Measurement Policy*, 1 (2007).
[317] FQS-I, *Uncertainty of Measurement*, FRAP-3, § 2.1 (2008).

a) WEIGHTED MEAN

i. "The weighted mean computation attaches more weight to those groups of measurements that are more precise whereas the arithmetic mean attaches equal weight to all measurements."[318]

1) "When there is significant variability in an analytical method and the known concentration is the objective, then a weighted mean computation is probably more appropriate."[319]

2) "When the solution measurements are made by different individuals and on different days, the simple arithmetic mean may not be the best estimate of the true solution value. A weighted mean may be a more appropriate estimate of the true concentration…It would seem that a weighted mean provides a better estimate of the true simulator solution value and should be employed for those cases in which significant inter-operator or interday variability exists. At the very least, the weighted mean should be compared to the arithmetic mean to determine if significant differences exist."[320]

ii. "Weighting data…is not some method of manipulating the result to make it appear more acceptable, *it is the correct statistical treatment for heteroscedastic data*."[321]

ii. UNCERTAINTY:

a) "Accounting for and documenting…measurement uncertainty is an important element of quality control…Forensic scientists, indeed all of those involved in the legal application of measurements, should appreciate its importance for establishing fitness-for-purpose."[322]

b) "Reliable analytical measurement expected to be forensically acceptable is far from trivial. Many elements converge as part of a well designed 'measurement algorithm' to produce results capable of being presented with confidence in a forensic context…Each component must be carefully considered regarding its contribution both to the confidence and uncertainty in the final result. Not only do the various elements help to ensure reliability but their individual characteristics also propagate uncertainty to the final result…The total magnitude of error, however, can be quantified in the final results to ensure acceptable limits."[323]

[318] Gullberg, *Using a Weighted Mean to Compute the Values of Simulator Solution Standards*, 14(3) J. ANAL. TOXICOL. 196-8 (1990).
[319] Gullberg, *Using a Weighted Mean to Compute the Values of Simulator Solution Standards*, 14(3) J. ANAL. TOXICOL. 196-8 (1990).
[320] Gullberg, *Using a Weighted Mean to Compute the Values of Simulator Solution Standards*, 14(3) J. ANAL. TOXICOL. 196-8 (1990).
[321] Flanagan, *Fundamentals of Analytical Toxicology* 370 (Wiley 2007).
[322] Gullberg, *Estimating the measurement uncertainty in forensic breath-alcohol analysis*, 11 ACCRED. QUAL. ASSUR. 562, 568 (2006).
[323] Gullberg, *Methodology and Quality Assurance in Forensic Breath Alcohol Analysis*, 12 For. Sci. Rev. 49, 50 (2000).

c) Forensic labs must "[c]onstruct and document an appropriate measurement uncertainty budget, identifying and listing all potential sources of uncertainty."[324]

d) "Records must be maintained to describe the process used to develop the estimation of uncertainty. These records must include the elements of the [uncertainty] budget, data gathered, calculations to arrive at the estimate, and the estimated uncertainty associated with the measurement method."[325]

iii. THE COVERAGE INTERVAL:

a) Estimates of uncertainty in forensic measurements are typically determined by computing expanded uncertainties and subsequent coverage intervals providing a desired level of confidence.[326]

b) The generally accepted coverage factor is $\lambda = 2$ or 3, yielding a level of confidence is 95% - 99%.[327]

$$Y = y \pm U$$
$$= y \pm \lambda \mu_c$$

iv. THE SAFETY MARGIN

a) A "valid approach is to make a deduction for uncertainty before the final result is reported to the court."[328] In this method, a result "is expressed as 'a minimum of' or 'not less than' a stated value, where an allowance has been made for the associated uncertainty of measurement. The allowance made for uncertainty is frequently in excess of the actual uncertainty."[329]

$$Y \geq y - s$$

d. FORENSIC DETERMINATION OF UNRELIABILITY

[324] ASCLD/LAB – International, *ESTIMATING UNCERTAINTY of MEASUREMENT POLICY*, 2 (2007).
[325] ASCLD/LAB – International, *ESTIMATING UNCERTAINTY of MEASUREMENT POLICY*, 3 (2007).
[326] Gullberg, *Estimating the measurement uncertainty in forensic breath-alcohol analysis*, 11 ACCRED. QUAL. ASSUR. 562, 568 (2006); Gullberg, *Breath Alcohol Measurement Variability Associated with Different Instrumentation and Protocols*, 131(1) FOR. SCI. INT. 30, 30 (2003).
[327] Jones, *Dealing with Uncertainty in Chemical Measurements*, 14(1) NEWSL. OF THE INT. ASSOC. FOR CHEM. TEST. 10 (2003); SWGDRUG, *Recommendations (Minimum Standards)*, §§ 4.3.2, 5.2.2 (2008); Gullberg, *Estimating the measurement uncertainty in forensic breath-alcohol analysis*, 11 ACCRED. QUAL. ASSUR. 562 (2006); Gullberg, *Breath Alcohol Measurement Variability Associated with Different Instrumentation and Protocols*, 131(1) FOR. SCI. INT. 30 (2003).
[328] Jones, *Dealing with Uncertainty in Chemical Measurements*, 14(1) NEWSL. OF THE INT. ASSOC. FOR CHEM. TEST. 10 (2003); Gullberg, *Estimating the measurement uncertainty in forensic breath-alcohol analysis*, 11 ACCRED. QUAL. ASSUR. 562, 567 (2006); Carpenter, *Breath Temperature: An Alabama Perspective*, 9(2) NEWSL. OF THE INT. ASSOC. FOR CHEM. TEST., 16, 17 (1998).
[329] Treble, *Analytical measurement and the law* 20 VAM BULLETIN 3, 4 (1999).

i. "Understanding the variables inherent in a measurement system that deals with [qualitative characteristics] is fundamental to a good quality assurance program."[330]

ii. FREQUENTIST METHODS: When engaging in qualitative analysis, "the paradigm of yes/no conclusions is useful for describing and quantifying the accuracy with which forensic science disciplines can provide answers. In such situations, results from analyses for which the truth is known can be classified in a two-way table."[331]

	Test Result A	Test Result ¬A	
Condition A	True Positive N_{TP}	False Negative (Type I error) N_{FN}	$N_{TP} + N_{FN}$
Condition ¬A	False Positive (Type II error) N_{FP}	True Negative N_{TN}	$N_{FP} + N_{TN}$
	$N_{TP} + N_{FP}$	$N_{FN} + N_{TN}$	N

a) As is the case generally, the forensic scientists can employ sensitivity, specificity, positive predictive value and negative predictive value as quantitative measures of unreliability.[332]

b) "[E]rrors and corresponding error rates can have more complex sources than can be accommodated within the simple framework presented above. For example, in the case of DNA analysis, a declaration that two samples match can be erroneous in at least two ways: The two samples might actually come from different individuals whose DNA appears to be the same within the discriminatory capability of the tests, or two different DNA profiles could be mistakenly determined to be matching. The probability of the former error is typically very low, while the probability of a false positive (different profiles wrongly determined to be matching) may be considerably higher. Both sources of error need to be explored and quantified in order to arrive at reliable error rate estimates for DNA analysis."[333]

iii. BAYES THEOREM: "Bayesian estimates are particularly widely used in evaluating forensic evidence, for example DNA matching or blood group matching."[334]

a) LIKELIHOOD RATIOS: The likelihood ratio "approach is now firmly established as a theoretical framework for any forensic discipline."[335] Since the likelihood ratio is

[330] Reeder, *Impact of DNA Typing on Standards and Practice in the Forensic Community* 123 ARCH. PATH. LAB. MED. 1063, 1064 (1999).
[331] NAS, *Strengthening Forensic Science in the United States: A Path Forward*, 4-6 (2009).
[332] NAS, *Strengthening Forensic Science in the United States: A Path Forward*, 4-7 (2009).
[333] NAS, *Strengthening Forensic Science in the United States: A Path Forward*, 4-8 – 4-9 (2009).
[334] Ellison, *Uncertainties in qualitative testing and analysis*, 5 ACCRED. QUAL. ASSUR. 346, 346 (2000); Ellison, *Quantifying uncertainty in qualitative analysis* 123 ANALYST 1155, 1157 (1998).
[335] Gonzalez-Rodriguez, *Biometric Identification in Forensic Cases According to the Bayesian Approach* in *Biometric Authentication* 177, 179 (Springer-Verlag 2002); Gonzalez-Rodriguez, *Emulating DNA: Rigorous Quantification of Evidential Weight in Transparent and Testable Forensic Speaker Recognition* 15(7) IEEE TRANS. AUDIO, SPEECH, AND LANG. PROCESSING 2104, 2104 (2007); Taroni,

a measure of the impact of the evidence on the initial hypothesis H, in a prosecution that the defendant is guilty/not guilty, it provides a quantitative measure of the relevance and weight of the evidence.[336]

b) POSTERIOR PROBABILITY:[337]

c) BAYESIAN NETWORKS: "In a Bayesian network, probability is associated with graph theory. Bayesian networks are a mathematically and statistically rigorous technique for representing and evaluating dependencies and influences among variables considered relevant for a particular inferential problem. Several authors have pointed out the utility of Bayesian networks for handling uncertainties associated with the evaluation of evidence in forensic science."[338]

3. REPORTING FORENSIC RESULTS

a. GENERAL CONSIDERATIONS

i. "Results of scientific measurements are compelling to those untrained in numerical or analytical issues while many believe that all numerical results possess absolute certainty. The professional expert witness, however, must present numerical information accompanied by their limitation and avoid conveying the "illusion of certainty". The misuse and misleading application of statistics, designed to convey an unjustified interpretation, must also be considered unethical. Doubt and uncertainty should be respectable concepts in the forensic sciences. While fitness-for-purpose can and should certainly be established, assumptions and uncertainty in breath alcohol analysis must be acknowledged."[339]

ii. "Communicating analytical results occurs during the post-analytical stage of a complete measurement process. No important measurement process is complete until the results have been clearly communicated to and understood by the appropriate decision maker. Forensic measurements are made for important reasons. People, often unfamiliar with analytical concepts, will be making important decisions based on these results. Part of the forensic toxicologist's responsibility is to communicate the best

Two Items of Evidence, No Putative Source An Inference Problem in Forensic Intelligence 51(5) J. FOR. SCI. 1350, 1351 (2006); Aitken, *Evaluation of trace evidence for three-level multivariate data with the use of graphical models* 50 COMP. STAT. DATA ANAL. 2571 (2006); Thompson, *How the Probability of a False Positive Affects the Value of DNA Evidence* 48(1) J. FOR. SCI. 47 (2003); Aitken, *Statistical Techniques and Their Role in Evidence Interpretation* in *Forensic Medicine: Clinical and Pathological Aspects* 755 (Greenwich Medical Media Ltd. 2002); Stockmarr, *Likelihood Ratios for Evaluating DNA Evidence When the Suspect is Found Through a Database Search* 55 BIOMETRICS 671 (1999).

[336] Stockmarr, *Likelihood Ratios for Evaluating DNA Evidence When the Suspect is Found Through a Database Search* 55 BIOMETRICS 671 (1999).

[337] Meester, *Why the Effect of Prior Odds Should Accompany the Likelihood Ratio When Reporting DNA Evidence* 3 LAW, PROB. AND RISK 51 (2004).

[338] Taroni, *Two Items of Evidence, No Putative Source An Inference Problem in Forensic Intelligence* 51(5) J. FOR. SCI. 1350, 1351 (2006); Dawid, *Object-oriented Bayesian networks for complex forensic DNA profiling problems* 169(2) FOR. SCI. INT. 195 (2007); Taroni, *Bayesian Networks and Probabilistic Inference in Forensic Science* (Wiley 2006); Bianchi, *Forensic DNA and bioinformatics* 8(2) BRIEFINGS IN BIOINFORMATICS 117 (2007).

[339] Gullberg, *Professional and Ethical Considerations in Forensic Breath Alcohol Testing Programs* 5(1) J. ALC. TEST. ALLIANCE 22, 25 (2006).

measurement estimate along with its uncertainty. Insufficient communication and interpretation of measurement results can introduce more uncertainty than the analytical process itself. The best instrumentation along with the most credible protocols ensuring the highest possible quality control will not compensate for the unclear and insufficient communication of measurement results and their significance."[340]

iii. "The terminology used in reporting and testifying about the results of forensic science investigations must be standardized."[341]

iv. "Calculations and data transfers which do not form part of a validated electronic process should be checked, preferably by a second person."[342]

v. "Before results are reported, each batch of analytical data should be reviewed by scientific personnel who are experienced with the analytical protocols used in the laboratory. At a minimum this review should include:… validity of analytical data (e.g., shape and signal-to-noise ratio of chromatographic peak) and calculations [and] quality control data."[343]

vi. "It is recognized that for a variety of reasons occasional analytical results will be outliers; that is, analytical values which deviate significantly and spuriously from the true value."[344]

a) If an outlier is suspected then it can be investigated utilizing Grubb's test.[345]

b. RESULT = MEASUREMENT + UNCERTAINTY

i. "[F]orensic test results must be validated and verified before they are presented to the court."[346]

ii. "If systematic error does exist this must be added or subtracted from the mean result."[347]

iii. When the result of a forensic measurement is reported simply as "'a number,' it does not reflect the accuracy of the measurement and cannot be properly interpreted." [348] "Estimating and reporting measurement uncertainty with the number completes the

[340] Gullberg, *Statistical Applications in Forensic Toxicology*, Medical-Legal Aspects of Alcohol, p. 457, 504 (James Garriott ed., 5th ed. 2009).
[341] NAS, *Strengthening Forensic Science in the United States: A Path Forward*, S-15 (2009).
[342] ILAC, *Guidelines for Forensic Science Laboratories*, ILAC-G19 § 4.12.2.1(e) (2002).
[343] SOFT/AAFS, *Forensic Toxicology Laboratory Guidelines*, § 10.1 (2006).
[344] SOFT/AAFS, *Forensic Toxicology Laboratory Guidelines*, § 8.3.9 (2006).
[345] Flanagan, *Fundamentals of Analytical Toxicology* 385 (Wiley 2007).
[346] Godowsky, *Quality Assurance in forensic Laboratories*, 4(6) EV. TECH. MAG. 36, 36 (2006).
[347] Jones, *Dealing with Uncertainty in Chemical Measurements*, 14(1) NEWSL. INT. ASSOC. CHEM. TEST. 10 (2003).
[348] Bono, *ISO/IEC 17025:2005: Section 5.4.6: Estimation of Uncertainty – Is Anyone Certain What This Means?* p.7, Presentation at the 61st Annual Meeting of the American Academy of Forensic Sciences (2/17/2009).

picture and allows us to properly use the result to make reliable and defensible decisions."[349]

iv. "Clear and sufficient communication of measurement results begins with adequate printed documentation. Measurement results and associated information read by decision makers should be clear, thorough and self-explanatory. The results must display…the associated uncertainty of the results. The uncertainty estimate can take the form of a…expanded uncertainty or a confidence interval…whenever possible, a numerical assessment of uncertainty should be provided."[350]

v. "All results for every forensic science method should indicate the uncertainty in the measurements that are made."[351]

vi. "Forensic reports, and any courtroom testimony stemming from them, must include clear characterizations of the limitations of the analyses, including measures of uncertainty in reported results and associated estimated probabilities where possible."[352]

vii. "[C]onsidering or not the uncertainty of a critical result can make the difference between acquittal and a guilty sentence."[353]

viii. COVERAGE INTERVAL

a) "Computing expanded uncertainties and subsequent confidence intervals for quantitative forensic evidence provides the court with relevant information for determining appropriate evidentiary weight."[354]

i. "For example, methods for measuring the level of blood alcohol in an individual or methods for measuring the heroin content of a sample can do so only within a confidence interval of possible values."[355]

ii. "An urgent need exists to report results of forensic alcohol analysis as a range of values, that is as a confidence statement."[356]

b) Forensic results need to be reported, along with a coverage interval that has a high

[349] Bono, *ISO/IEC 17025:2005: Section 5.4.6: Estimation of Uncertainty – Is Anyone Certain What This Means?* p.7, Presentation at the 61st Annual Meeting of the American Academy of Forensic Sciences (2/17/2009).
[350] Gullberg, *Statistical Applications in Forensic Toxicology*, Medical-Legal Aspects of Alcohol, p. 457, 504-505 (James Garriott ed., 5th ed. 2009).
[351] NAS, *Strengthening Forensic Science in the United States: A Path Forward*, 6-1 (2009); Gullberg, *Breath Alcohol Measurement Variability Associated with Different Instrumentation and Protocols*, 131(1) FOR. SCI. INT. 30, 30 (2003).
[352] NAS, *Strengthening Forensic Science in the United States: A Path Forward*, S-16, 6-3 (2009).
[353] Bich, *Interdependence between measurement uncertainty and metrological traceability* ACCRED. QUAL. ASSUR. (IN PRESS - 2009).
[354] Gullberg, *Estimating the measurement uncertainty in forensic breath-alcohol analysis*, 11 ACCRED. QUAL. ASSUR. 562, 568 (2006).
[355] NAS, *Strengthening Forensic Science in the United States: A Path Forward*, 4-5 (2009); Kristiansen, *An Uncertainty Budget for the Measurement of Ethanol in Blood by Headspace Gas Chromatography*, 28(6) J. ANAL. TOX. 456 (2004); Jones, *Dealing with Uncertainty in Chemical Measurements*, 14(1) NEWSL. OF THE INT. ASSOC. FOR CHEM. TEST. 6 (2003).
[356] A.W. Jones, Ph.D, *Dealing with Uncertainty in Chemical Measurements*, 14(1) Newsletter of the International Association for Chemical Testing 6, 7 (2003).

probability of containing the true value of the measurand, that is "the mean plus or minus two standard deviations."[357]

ix. SAFETY MARGIN

a) Jurisdictions in both the U.S. and Europe utilize this method expressing a result as a minimum value with a stated level of confidence after subtracting the expanded uncertainty from the result.[358]

c. RESULT OF QUALITATIVE FORENSIC TEST OBSERVATION:

i. "Forensic scientists are required to qualify and, where possible, quantify their states of knowledge and to be consultants in the assessment of uncertainties associated with the inferences that may be drawn from forensic evidence."[359]

ii. The results of identification evidence should be limited to the reporting of a likelihood ratio. This gives the court and the trier of fact information concerning the relative strength and impact of the evidence on the determination to be made so that it may be assigned appropriate weight.[360]

F. SCIENTIFIC FORENSIC STANDARDS

1. UTILITY

a. Forensic science needs standards governing "protocols for forensic examinations, methods, and practices" to ensure application of best practices in "measurement, validation, reliability…and proficiency testing in forensic science."[361]

b. "Forensic science stakeholders need to be assured that the profession is following standard methodology, so that the stakeholders have a way of judging whether the forensic science results are accurate, reliable, or meaningful in the context of the case they are dealing with."[362]

c. "Standard practices, specifications, and test methods make it possible for business to be conducted in a workmanlike manner with all participants having confidence in the validity and reliability of the measurements and analyses involved. In this regard, forensic science should follow the example set by the rest of the business and scientific world."[363]

[357] NAS, *Strengthening Forensic Science in the United States: A Path Forward*, 4-6 (2009).
[358] Gullberg, *Estimating the measurement uncertainty in forensic breath-alcohol analysis*, 11 ACCRED. QUAL. ASSUR. 562, 567 (2006); Jones, *Dealing with Uncertainty in Chemical Measurements*, 14(1) NEWSL. OF THE INT. ASSOC. FOR CHEM. TEST. 10 (2003); Carpenter, *Breath Temperature: An Alabama Perspective*, 9(2) NEWSL. OF THE INT. ASSOC. FOR CHEM. TEST., 16, 17 (1998).
[359] Taroni, *Bayesian Networks and Probabilistic Inference in Forensic Science*, Preface (Wiley 2006).
[360] Champod, *Identification and Individualization* in *Wiley Encyclopedia of Forensic Science* (*in press* 2009).
[361] NAS, *Strengthening Forensic Science in the United States: A Path Forward*, 7-18 (2009).
[362] Lentini, *Forensic Science Standards: Where They Come From and How They Are Used*, 1 FOR. SCI. POL. MGMT. 10, 16 (2009).
[363] Lentini, *Forensic Science Standards: Where They Come From and How They Are Used*, 1 FOR. SCI. POL. MGMT. 10, 16 (2009).

d. Standardization in forensic science reduces confusion, eliminates causes of error and makes it possible for independent evaluation of results.[364]

e. "More standards to support the accuracy of testing and thorough comprehensive reviews of forensic laboratories should be embraced by all levels of the scientific community."[365]

2. GOLD STANDARDS: Standards promulgated by ISO, NIST and ASTM are nearly universally recognized throughout the forensics community.[366]

a. ISO:

i. "Any laboratory seeking ASCLD/LAB–International accreditation must demonstrate conformance to the requirements in ISO/IEC 17025."[367] "Conforming to the numbered requirements in [ISO 17025] is mandatory."[368]

b. NIST

i. "Scientific research at NIST starts with understanding the fundamentals of science, from which standards are created. These standards are the focal point of the forensic science program at the Office of Law Enforcement Standards…the end result is a standard that provides the necessary basis by which forensic analysts provide the scientific results that meet judicial acceptability."[369]

c. ASTM

i. ASTM Committee E30 on Forensic Sciences was founded by members of the American Academy of Forensic Sciences and currently maintains more than fifty published forensic science standards. "Most public forensic science laboratories in the United States have at least one member participating in the ASTM process."[370]

[364] Lentini, *Forensic Science Standards: Where They Come From and How They Are Used*, 1 FOR. SCI. POL. MGMT. 10, 16 (2009); King, *Chemical measurement and the law: metrology and quality issues*, 6 ACCRED. QUAL. ASSUR. 236, 238-9 (2001).
[365] FQS-I, *Position Statement Regarding NAS Report* (2009).
[366] Lentini, *Forensic Science Standards: Where They Come From and How They Are Used*, 1 FOR. SCI. POL. MGMT. 10 (2009); FQS-I, *Forensic Requirements for Accreditation*, FRA-1 (2008); FSAB, *Standards For Accrediting Forensic Specialty Certification Boards*, 1 (2004); SOFT/AAFS, *Forensic Toxicology Laboratory Guidelines*, § 9.3.1 (2006); OIML, *Breath Alcohol Analysers*, 2 (2006); NAS, *Strengthening Forensic Science in the United States: A Path Forward*, 7-4 – 7-7, 7-18 (2009); The Crime Lab Report, *How the Profession Was Revolutionized By Standards And Controls* (2007); Gonzalez-Rodriguez, *Emulating DNA: Rigorous Quantification of Evidential Weight in Transparent and Testable Forensic Speaker Recognition* 15(7) IEEE TRANS. AUDIO, SPEECH, AND LANG. PROCESSING 2104, 2104 (2007); Reeder, *Impact of DNA Typing on Standards and Practice in the Forensic Community* 123 ARCH. PATH. LAB. MED. 1063 (1999).
[367] ASCLD/LAB – International, *Lab International Accreditation Program*, 2 (2006); ASCLD/LAB – International, *Breath Alcohol Calibration Accreditation Program*, 3 (2008).
[368] ASCLD/LAB – International, *Lab International Accreditation Program*, 2 (2006).
[369] NIST, *Office of Law Enforcement Standards: Programs, Activities and Accomplishments*, NISTIR 7366, 26 (2007).
[370] Lentini, *Forensic Science Standards: Where They Come From and How They Are Used*, 1 FOR. SCI. POL. MGMT. 10, 12-15 (2009).

3. "Appropriate [forensic] standards must be coupled with effective systems of accreditation and/or certification that include strong enforcement mechanisms and sanctions."[371]

VI. SPECIAL TOPICS

A. DNA

1. STATISTICAL APPROACH TO PRESENTATION OF RESULTS[372]

a. "The way in which statistical DNA evidence is presented to legal decision makers can have a profound impact on the persuasiveness of that evidence."[373]

b. Research Findings – what to look out for: [374]

i. "…in general, people attach less weight to the statistical evidence than would seem appropriate."

ii. "…jurors had trouble aggregating a 1 in 1 billion DNA match statistic with laboratory error rate statistics."

iii. "…jurors underestimated the probative value of DNA evidence relative to Bayesian norms."

iv. "…laypeople are not intuitive Bayesians in cases involving DNA statistics, and they may not assess the probative value of a DNA match in clear and consistent ways."

c. Research Findings – what might be useful:[375]

i. "…laypeople tend to be more impressed with DNA statistics when they are presented as likelihood ratios rather than as frequencies."

ii. Important psychological difference between:[376]

a) Recognizing possibility that DNA match arose by coincidence.

i. "[S]mall, abstract chance may be treated as essentially zero."

b) Realizing that coincidental matches exist and are plentiful.

i. "[E]xemplars transform mere statistical possibility into imagery that is more compelling."

---

[371] NAS, *Strengthening Forensic Science in the United States: A Path Forward*, 7-10 (2009).
[372] Koehler, *When Are People Persuaded By DNA Match Statistics?* 25(5) LAW AND HUMAN BEHAVIOR 493 (2001).
[373] Koehler, *When Are People Persuaded By DNA Match Statistics?* 25(5) LAW AND HUMAN BEHAVIOR 493, 493 (2001).
[374] Koehler, *When Are People Persuaded By DNA Match Statistics?* 25(5) LAW AND HUMAN BEHAVIOR 493, 494-495 (2001).
[375] Koehler, *When Are People Persuaded By DNA Match Statistics?* 25(5) LAW AND HUMAN BEHAVIOR 493, 495 (2001).
[376] Koehler, *When Are People Persuaded By DNA Match Statistics?* 25(5) LAW AND HUMAN BEHAVIOR 493, 508-509 (2001).

d. Strategic and policy implications:[377]

i. "[P]rosecution should present DNA statistics in a single-target, probability frame format. This presentation makes it difficult to take seriously the possibility that the match is merely coincidental."

ii. "The defense should favor a multi-target, frequency frame format in cases where exemplar generation seems reasonable (i.e., where the incidence rate is not smaller than the reference class jurors will most likely use)."

iii. "Judges and legislators may also find this research useful when considering standards for presenting scientific and statistical evidence in court. For example, judicial instructions might be formulated that acknowledge that there are different ways of presenting the same statistical information."

2. UNCERTAINTY/UNRELIABILITY

a. "[I]n the case of DNA analysis, a declaration that two samples match can be erroneous in at least two ways: The two samples might actually come from different individuals whose DNA appears to be the same within the discriminatory capability of the tests, or two different DNA profiles could be mistakenly determined to be matching. The probability of the former error is typically very low, while the probability of a false positive (different profiles wrongly determined to be matching) may be considerably higher. Both sources of error need to be explored and quantified in order to arrive at reliable error rate estimates for DNA analysis."[378]

b. "When evaluating the strength of DNA evidence for proving that two samples have a common source, one must consider two factors. One factor is the probability of a coincidental match (sometimes called the random match probability). A coincidental match occurs when two different people have the same DNA profile. The second factor is the probability of a false positive. A false positive (as we use that term here) occurs when a laboratory erroneously reports a DNA match between two samples that actually have different profiles…Either a coincidental match or a false positive could cause a laboratory to report a DNA match between samples from different people. Consequently, one must consider both the random match probability and the false positive probability in order to make a fair evaluation of DNA evidence."[379]

[377] Koehler, *When Are People Persuaded By DNA Match Statistics?* 25(5) LAW AND HUMAN BEHAVIOR 493, 509 (2001).
[378] NAS, *Strengthening Forensic Science in the United States: A Path Forward*, 4-8 – 4-9 (2009).
[379] Thompson, *How the Probability of a False Positive Affects the Value of DNA Evidence* 48(1) J. FOR. SCI. 47 (2003).

i. Sources of false positives include:[380]

a) Error in the collection or handling of samples;

b) Accidently switching reference samples of victim and defendant;

c) Misinterpretation of results;

d) Equipment effects;

ii. "Ignoring or underestimating the potential for a false positive can lead to serious errors of interpretation."[381]

B. BLOOD ALCOHOL TESTING

1. METHODS

a. "[M]ust be a recognized method having the requisite reliability, and it must be accompanied by adequate quality assurance procedures."[382]

i. Gas chromatography alone or in conjunction with mass spectrometry is a widely accepted method for analysis of both alcohol and most drugs in blood and is the method treated herein.[383]

2. QUALITY ASSURANCE

a. "[I]t is important to integrate into the laboratory's good laboratory practices, as a minimum…establishment and validation of calibrations [and] checks on linearity, and other analysis instructions provided by the applicable manufacturer(s) of the instrument(s) and commercial reagents utilized. Further, the ongoing mandates on use of controls and other good laboratory practices promulgated by…applicable authority should be recognized and complied with when applicable."[384]

[380] Thompson, *How the Probability of a False Positive Affects the Value of DNA Evidence* 48(1) J. FOR. SCI. 47 (2003).
[381] Thompson, *How the Probability of a False Positive Affects the Value of DNA Evidence* 48(1) J. FOR. SCI. 47, 56 (2003).
[382] NCCLS, *Blood Alcohol Testing in the Clinical Laboratory; Approved Guideline*, T/DM6-A § 1.4.6 (1997).
[383] Westphal, *Development of a validated method for the simultaneous determination of amphetamine, methamphetamine and methylenedioxyamphetamines (MDA, MDMA, MDEA) in serum by GC-MS after derivatisation with perfluorooctanoyl chloride* 12 ACCRED. QUAL. ASSUR. 335 (2007); SOFT/AAFS, *Forensic Toxicology Laboratory Guidelines*, § 8 (2006); Moeller, *Drugs of Abuse Monitoring in Blood for Control of Driving Under the Influence of Drugs* 24 THER. DRUG MON. 210 (2002); Moeller, *Determination of drugs of abuse in blood* 713 J. CHROM. B, 91 (1998); NCCLS, *Blood Alcohol Testing in the Clinical Laboratory; Approved Guideline*, T/DM6-A § 4.1 (1997). "Urine alcohol concentration is not a good indicator of intoxication. Urine alcohol concentration is dependent upon number of hours elapsed since last voiding, fluid intake, and number of alcoholic drinks consumed." Mayo Clinic, *2008 Drug Testing An Overview of Mayo Clinic Tests Designed for Detecting Drug Abuse* (2008).
[384] NCCLS, *Blood Alcohol Testing in the Clinical Laboratory; Approved Guideline*, T/DM6-A § 5 (1997).

b. CALIBRATION

i. "The calibrators should be selected to represent critical concentrations, which span the clinically and forensically relevant alcohol concentrations and include the upper limit of linearity of the analysis. These calibrators will bracket the majority of positive results and can be used to demonstrate linearity."[385]

ii. "Aqueous Standard Reference Materials containing ethanol are available from the National Institute of Standards and Technology (SRM 1828a)."[386]

c. TESTING

i. "Every alcohol analysis or batch of analyses performed by GC methods should begin with the analysis of at least one, and preferably two or more different calibrators together with an alcohol-free 'blank,' because the operating parameters and calibration of GC instruments vary with each startup and can also drift during prolonged operation."[387]

ii. "Every analysis or batch of analyses should be accompanied by the analysis of negative and positive controls."[388]

iii. "Because of the variability of instrument parameters and calibration with each startup, and the tendency of these factors to drift during prolonged instrument operation, at least every tenth specimen should be a control when multiple, sequential analyses are conducted."[389]

iv. "With each batch of specimens, whether a single specimen or multiple ones, controls would be carried through the procedure in parallel with the unknowns. It is suggested that each batch of specimens include at least 10% controls. The controls must include one positive and one negative control. For qualitative assays positive and negative controls, acceptable results may simply be positive or negative, respectively."[390]

[385] NCCLS, *Blood Alcohol Testing in the Clinical Laboratory; Approved Guideline*, T/DM6-A § 5.1 (1997); SOFT/AAFS, *Forensic Toxicology Laboratory Guidelines*, § 8.3.6 (2006); Flanagan, *Fundamentals of Analytical Toxicology* 357 (Wiley 2007).
[386] NCCLS, *Blood Alcohol Testing in the Clinical Laboratory; Approved Guideline*, T/DM6-A § 5.1 (1997); SOFT/AAFS, *Forensic Toxicology Laboratory Guidelines*, § 9.3.1 (2006).
[387] NCCLS, *Blood Alcohol Testing in the Clinical Laboratory; Approved Guideline*, T/DM6-A § 5.1 (1997).
[388] NCCLS, *Blood Alcohol Testing in the Clinical Laboratory; Approved Guideline*, T/DM6-A § 5.2 (1997).
[389] NCCLS, *Blood Alcohol Testing in the Clinical Laboratory; Approved Guideline*, T/DM6-A § 5.2 (1997).
[390] SOFT/AAFS, *Forensic Toxicology Laboratory Guidelines*, § 9.2.1 (2006).

d. PROFICIENCY TESTING

i. "Forensic toxicology laboratories should participate in an external proficiency testing program which includes, at a minimum, samples for alcohol in blood or serum, and for drugs in at least one type of specimen, representative of that typically analyzed by the laboratory (e.g. whole blood or serum for a postmortem toxicology laboratory). The program should realistically monitor the laboratory's quantitative capability."[391]

3. STANDARDS

a. Failure to adhere to standards of acceptable quality of analysis leads to large variations in results of analysis on control samples between different laboratories.[392]

b. "The errors that can occur during the collection and handling of blood specimens are potentially numerous (e.g., inaccurate identification of specimens, specimen hemolysis, the improper handling of anticoagulants, the formation of hematomas, hemoconcentration). Standards for venipuncture can reduce or alleviate many of these errors in much the same way that quality control standards have reduced errors within the laboratory."[393]

4. UNCERTAINTY

a. SOURCES OF UNCERTAINTY

i. Blood Draw: "Factors that Effect Laboratory Values…Major causes of 'laboratory error' can be related to nonanalytical factors such as specimen collection, handling, and transport."[394] "The errors that can occur during the collection and handling of blood specimens are potentially numerous"[395]

a) "Nonbiological factors—such as patient misidentification…contribute to the total 'laboratory error.'"[396]

b) "[B]iological factors—such as patient posture and the time a specimen is drawn, [] contribute to the total 'laboratory error.'"[397]

---

[391] SOFT/AAFS, *Forensic Toxicology Laboratory Guidelines*, § 9.1.9 (2006).
[392] Falkensson, *Hospital alcohol analyses not completely reliable External quality control needed at least every year* 87 LÄKARTIDNINGEN 470 (1990).
[393] NCCLS, *Procedures for the Collection of Diagnostic Blood Specimens by Venipuncture; Approved Standard—Fifth Edition*, H3-A5, vii (2003).
[394] NCCLS, *Procedures for the Collection of Diagnostic Blood Specimens by Venipuncture; Approved Standard—Fifth Edition*, H3-A5, § 5 (2003).
[395] NCCLS, *Procedures for the Collection of Diagnostic Blood Specimens by Venipuncture; Approved Standard—Fifth Edition*, H3-A5, vii (2003); Ashavaid, *Influence of Blood Specimen Collection Method on Various Preanalytical Sample Quality Indicators* 23(2) INDIAN J. CLIN. BIOCHEM. 144(2008); Ernst, *Preanalytical Errors that Occur During Specimen Collection*, Articles in Phlebotomy Center for Phlebotomy Education (2007).
[396] NCCLS, *Procedures for the Collection of Diagnostic Blood Specimens by Venipuncture; Approved Standard—Fifth Edition*, H3-A5, § 5 (2003).
[397] NCCLS, *Procedures for the Collection of Diagnostic Blood Specimens by Venipuncture; Approved Standard—Fifth Edition*, H3-A5, § 5 (2003).

c) Contamination

i. "If the venipuncture proves difficult and the vein must be touched again to draw blood, the site should be cleansed again."[398]

ii. "When drawing a blood specimen for alcohol testing, a nonalcohol-based cleanser should be used to cleanse the venipuncture site."[399]

d) Hemolysis: "Alteration, dissolution, or destruction of red blood cells in such a manner that hemoglobin is liberated into the medium in which the cells are suspended."[400]

i. "To prevent hemolysis when performing a venipuncture, the phlebotomist should: After cleansing, allow the venipuncture site to air dry; Never draw blood through a hematoma; If using a syringe, make sure the needle is fitted securely on a syringe to avoid frothing; When using a syringe and needle, avoid drawing the plunger back too forcibly; Gently invert the blood collection tube to mix additive specimens as recommended by the manufacturer."[401]

ii. Specimen Handling and Storage effects: "There are many analytical factors that may invalidate what appears to be an otherwise valid measurement. If the sample is transported or stored incorrectly, then no matter how good the measurement system is, the result produced may well be meaningless. The result of these transportation and storage effects can either raise or lower the apparent concentration of the target analyte."[402]

a) "Storage effects can not only be reflected in the loss of an analyte. It is well known in forensic toxicology that metabolites of drugs can be converted back to the parent compound, elevating the apparent concentration and possibly indicating a drug overdose where one does not exist."[403] The types of effects can be categorized as:[404]

i. Primary effect: "[A]cts externally on the sample to alter the energy entering, or the environment of, the sample."

ii. Secondary effect: The action of something contained within the sample.

[398] NCCLS, *Procedures for the Collection of Diagnostic Blood Specimens by Venipuncture; Approved Standard—Fifth Edition*, H3-A5, § 8.8.3 (2003).
[399] NCCLS, *Procedures for the Collection of Diagnostic Blood Specimens by Venipuncture; Approved Standard—Fifth Edition*, H3-A5, § 11.2.1 (2003); NCCLS, *Blood Alcohol Testing in the Clinical Laboratory; Approved Guideline*, T/DM6-A § 2.3.3 (1997).
[400] *Stedman's Medical Dictionary for the Health Professions and Nursing* 701 (6th ed. 2008).
[401] NCCLS, *Procedures for the Collection of Diagnostic Blood Specimens by Venipuncture; Approved Standard—Fifth Edition*, H3-A5, § 10.3 (2003).
[402] Williams, *How do storage conditions affect your samples?* 20 VAM BULLETIN 22, 22 (1999); NCCLS, *Procedures for the Collection of Diagnostic Blood Specimens by Venipuncture; Approved Standard—Fifth Edition*, H3-A5, § 5 (2003).
[403] Williams, *How do storage conditions affect your samples?* 20 VAM BULLETIN 22, 22 (1999).
[404] Williams, *How do storage conditions affect your samples?* 20 VAM BULLETIN 22, 22 (1999).

iii. Analyte effect: Transformation of analyte due to inherent propertied of analyte.

b) Microorganisms: Secondary effect. "[I]t has been documented that changes produced by contaminating microorganisms can affect alcohol concentrations in blood specimens even in the presence of preservatives…various organisms isolated from contaminated blood specimens [are] capable of producing ethanol when inoculated into bank blood. *Candida albicans* [is] particularly active in this regard, producing significant quantities of alcohol even in the presence of sodium fluoride…investigators recommended that fluoride (10 mg/mL; 0.24mmol/ml) be used as a preservative and that care should be taken to assure that microbial organisms are not introduced into the specimens."[405]

i. "The time of collection is critical information which must be recorded and should appear on the report of results."[406]

ii. "For whole blood or plasma specimens…sodium fluoride (1.5 mg/mL of blood; 3.6Fmol/ml) [is] an appropriate [amount of] preservative for storage at 5°C of initially sterile blood specimens for up to 48 hours. Blood alcohol specimens stored at -20 EC or below are stable indefinitely. Specimens that are to be transported or mailed in an unrefrigerated condition, or stored for more than 48 hours should be preserved with higher concentrations of sodium fluoride (10 mg/mL of blood; 0.24mmol/mL)."[407]

iii. "To ensure complete dissolution of the fluoride in the blood, the closed container of blood should be gently inverted several times immediately following specimen collection."[408]

c) "The attribution of a single cause for the loss or production of analytes is probably not realistic. While the addition of a preservative may help the situation, or a reduction of the energy entering a system may reduce the problems, these actions may in themselves cause problems. Consequently, the most reliable way to ensure that the analytical results produced from an individual sample are as realistic as possible, the sample should be analysed as soon after collection as possible."[409]

iii. Type of Specimen: "[T]he alcohol concentration of whole blood is not identical to that of plasma or of serum…theoretical calculations, based on water content, and experimental data yield typical mean ratios of 1.12/1 to 1.18/1 in normal subjects for

[405] NCCLS, *Blood Alcohol Testing in the Clinical Laboratory; Approved Guideline*, T/DM6-A § 2.3.4 (1997); Dick, *Alcohol Loss Arising From Microbial Contamination of Drivers' Blood Specimens* 34 FOR. SCI. INT. 17 (1987); Blume, *The Effect of Micribial Contamination of the Blood Sample on the Determination of Ethanol Levels in Serum* 60 AM. J. CLIN. PATH. 700 (1973); Jones, *Salting-out effect of Sodium Floride and its Influence on the Analysis of Ethanol by Headspace Gas Chromatography*, 18 J. ANAL. TOX. 292 (1994).
[406] NCCLS, *Blood Alcohol Testing in the Clinical Laboratory; Approved Guideline*, T/DM6-A § 2.3.1 (1997).
[407] NCCLS, *Blood Alcohol Testing in the Clinical Laboratory; Approved Guideline*, T/DM6-A § 2.3.1 (1997); Chang, *The Effect of Temperature on the Formation of Ethanol by Candida Albicans in Blood* 34(1) J. FOR. SCI. 105 (1989).
[408] NCCLS, *Blood Alcohol Testing in the Clinical Laboratory; Approved Guideline*, T/DM6-A § 2.3.1 (1997).
[409] Williams, *How do storage conditions affect your samples?* 20 VAM BULLETIN 22, 24 (1999).

serum/whole blood alcohol concentrations, with typical experimental ranges of 1.05/1 to 1.25/1."[410]

iv. Physiological Factors: "Physiological factors that influence results include age, activity, bed rest, food ingestion, alcohol ingestion, menstrual cycle, obesity, oral contraceptives, posture, pregnancy, race, gender, smoking, and time of day. All biological phenomena exhibit rhythms, with the circadian rhythm (the change in a 24-hour period) being the most important to laboratory testing. Many factors with documented effects on laboratory values have been published."[411]

b. AN UNCERTAINTY BUDGET FOR THE MEASUREMENT OF ETHANOL IN BLOOD BY HEADSPACE GAS CHROMATOGRAPHY:[412]

i. Includes four sources of uncertainty:[413]

a) Analytical

b) Traceability

c) Density of blood (average)

d) Interindividual variation in blood water content

ii. Deliberately omitted sampling/collection, handling/storage and transportation contributions to measurement uncertainty for purposes of this analysis.[414]

a) "Ethanol levels in blood may change after sampling. Sodium fluoride is added to blood sampling vials to avoid biological production or consumption of ethanol."[415]

iii. Conclusion:

a) "When measuring fresh blood, the relative combined standard uncertainty is in the order of 1.6% in the middle of the concentration range (1.2–2.0 g/kg) and increases to approximately 5% at 0.2 g/kg. It also increases slightly in the range of 2.0 to 3.0 g/kg because of an increase in the analytical relative standard uncertainty."[416]

---

[410] NCCLS, *Blood Alcohol Testing in the Clinical Laboratory; Approved Guideline*, T/DM6-A § 2.2 (1997).
[411] NCCLS, *Procedures for the Collection of Diagnostic Blood Specimens by Venipuncture; Approved Standard—Fifth Edition*, H3-A5, § 5 (2003).
[412] Kristiansen, *An Uncertainty Budget for the Measurement of Ethanol in Blood by Headspace Gas Chromatography*, 28(6) J. ANAL. TOX. 456 (2004).
[413] Kristiansen, *An Uncertainty Budget for the Measurement of Ethanol in Blood by Headspace Gas Chromatography*, 28(6) J. ANAL. TOX. 456, 457 (2004).
[414] Kristiansen, *An Uncertainty Budget for the Measurement of Ethanol in Blood by Headspace Gas Chromatography*, 28(6) J. ANAL. TOX. 456, 462 (2004).
[415] Kristiansen, *An Uncertainty Budget for the Measurement of Ethanol in Blood by Headspace Gas Chromatography*, 28(6) J. ANAL. TOX. 456, 462 (2004).
[416] Kristiansen, *An Uncertainty Budget for the Measurement of Ethanol in Blood by Headspace Gas Chromatography*, 28(6) J. ANAL. TOX. 456, 463 (2004).

b) Assuming "a sample with a true concentration 2.00 g/kg [and] the legal limit is 2.00 g/kg, half of the measurement results will exceed this limit because of the uncertainty of the measurement. However, very few results exceed 2.1 g/kg, hence, a subtraction of 0.1 g/kg from the result is sufficient in most cases to avoid erroneously concluding that the limit is exceeded when in fact it is not."[417]

c) "[F]or fresh blood measurements, the probability of committing a type 1 error is less than 0.1% with a safety margin of 0.1 g/kg, at least up to a concentration level of 2.00 g/kg (assuming $n = 2$)…It should be emphasized that only the combined standard uncertainty should be used to establish such safety margins. The analytical uncertainty is a part of the combined standard uncertainty of measurement; hence, basing the safety margin on the analytical uncertainty alone will overestimate the safety provided by it."[418]

d) Conversion:[419] g/Kg → g/100ml

i. BAC g/100mL ≡ .1055 · BAC g/Kg

ii. Safety Margin = .1 g/Kg
= .0106 g/100ml

iv. EXTENSION: Assume the result compensated for by reduction of the safety margin determined is represented by BAC ≥ $R_{esult} - U$ (99% level of confidence). Since sampling/collection, handling/storage and transportation contributions to measurement uncertainty were deliberately omitted, their recognized effects decrease the level of confidence bestowed by the author's safety margin. If we wish to obtain a realistic safety margin with the same level of confidence, then we need to combine the contributions made by these other sources of uncertainty with the combined standard uncertainty found by the author which will yield a new safety margin $\varepsilon > U$.

[417] Kristiansen, *An Uncertainty Budget for the Measurement of Ethanol in Blood by Headspace Gas Chromatography*, 28(6) J. ANAL. TOX. 456, 463 (2004).
[418] Kristiansen, *An Uncertainty Budget for the Measurement of Ethanol in Blood by Headspace Gas Chromatography*, 28(6) J. ANAL. TOX. 456, 463 (2004).
[419] Brick, *Standardization of Alcohol Calculations in Research* 30(8) ALC. CLIN. EXP. RES. 1276, 1285 (2006); ; Jones, *Alcohol test at hospital not easily applicable for judicial purposes Conversion of ethanol concentration in plasma or serum to blood alcohol level* 105 LÄKARTIDNINGEN 367 (2008).

5. THE PROBLEM WITH CONCENTRATION

a. Different Concepts Employed as Concentration: [420]

Physical Concept	Symbol	Definition	Units	Chem. Term.
Amount/Substance concentration	c	n/V	mol/m^3, mol/L	
Mass concentration	ρ	m/V	kg/m^3, kg/L	% w/v
Mass fraction	ω_b	$m_b/\sum m_i$	%	% w/w
Volume Fraction	φ_b	$V_b/\sum V_i$	%	% v/v

b. In the field of blood alcohol testing, the term concentration does not have a unique physical meaning.[421] Reporting conventions include:[422]

BAC Concept Measure	0.08% w/v	0.10% w/v
$\rho \rightarrow$ g/dL	0.08	0.10
$\rho \rightarrow$ g/100 mL	0.08	0.10
$\omega_b \rightarrow$ mg/g (%)	0.76	0.95
$c \rightarrow$ mmol/L	17.3	21.7
$\rho \rightarrow$ mg/dL	80	100
$\rho \rightarrow$ g/L	0.80	1.0

c. Converting between results is not always simply a matter of translating between different units but sometimes between different physical concept entities altogether which can lead to ambiguity, confusion and greater quantitative uncertainty.[423] The metrologically sound practice would be to standardize units and computations so that results reported by different individuals could be readily compared and understood.[424]

---

[420] BIPM, *The International System of Units (SI)* §§ 2.2.1, 5.3.7 (8th ed. 2006); IUPAC, *Quantities, Units and Symbols in Physical Chemistry* 47-48 (3rd ed. 2007); *CRC Handbook of Chemistry and Physics* 2-8 (89th ed. 2008); Watson, *Pharmaceutical Analysis - A Textbook for Pharmacy Students and Pharmaceutical Chemists*, 17-20 (2nd ed. Elsevier 2005); Dybkaer, *The meaning of 'concentration'*12 ACCRED. QUAL. ASSUR. 661 (2007).

[421] Imobersteg, *Attacking and Defending Drunk Driving Cases* § 9.02 (2008); Brick, *Standardization of Alcohol Calculations in Research* 30(8) ALC. CLIN. EXP. RES. 1276 (2006); NCCLS, *Blood Alcohol Testing in the Clinical Laboratory; Approved Guideline*, T/DM6-A § 6.2 (1997).

[422] Imobersteg, *Attacking and Defending Drunk Driving Cases* § 9.02 (2008); Brick, *Standardization of Alcohol Calculations in Research* 30(8) ALC. CLIN. EXP. RES. 1276 (2006); NCCLS, *Blood Alcohol Testing in the Clinical Laboratory; Approved Guideline*, T/DM6-A § 6.2 (1997); Simel, *Blood Alcohol Measurements in the Emergency Department:Who Needs Them?* 78(11) AJPH 1478 (1988).

[423] Dybkaer, *The meaning of 'concentration'*12 ACCRED. QUAL. ASSUR. 661 (2007); Brick, *Standardization of Alcohol Calculations in Research* 30(8) ALC. CLIN. EXP. RES. 1276 (2006).

[424] Dybkaer, *The meaning of 'concentration'*12 ACCRED. QUAL. ASSUR. 661 (2007); Brick, *Standardization of Alcohol Calculations in Research* 30(8) ALC. CLIN. EXP. RES. 1276 (2006).

VII. FORENSIC METROLOGY AND THE LAW

A. GENERAL PRINCIPLES:

1. "In this age of science we must build legal foundations that are sound in science as well as in law."[425]

2. "The law should seek verdicts consistent with scientific reality…and it can achieve this goal only by requiring scientific evidence to conform to the standards and criteria to which scientists themselves adhere."[426]

3. "[I]n order to qualify as 'scientific knowledge,' an inference or assertion must be derived by the scientific method."[427]

B. DAUBERT,[428] FRYE[429] AND EVID. R. 702:

1. "In a case involving scientific evidence, *evidentiary reliability* will be based upon *scientific validity.*"[430] "The term 'scientific' implies a grounding in the methods and procedures of science."[431] It "draws its convincing force from some principle of science, mathematics and the like."[432] "[A] hypothesis that cannot be subject to the possibility of rejection by observation and experiment cannot be regarded as scientific."[433] "[I]ndeed this methodology distinguishes science from other fields of human inquiry."[434]

2. Under *Daubert*, courts must engage in "a preliminary assessment of whether the reasoning or methodology underlying the testimony is scientifically valid."[435] There is "an inherent limitation in the process of judicial evaluation of the reliability and validity of any scientific or technical evidence: the court…is limited in its ability to do so by the quantitative and qualitative nature of the evidence produced by the parties, whatever research the court itself may do, and any help it may derive from courts that have addressed the issue before it. This process unavoidably takes place on a continuum, and a court faced with the present task of deciding the admissibility of scientific evidence must exercise care to consider whether new developments or evidence require a reevaluation of the conclusions previously reached by courts that did not have the benefit of the more recent information. In short, neither science and technology may rest on past

---

[425] Justice Stephen Breyer in, *Reference Manual on Scientific Evidence* 4 – 8 (2nd ed. 2000).
[426] Black, *Evolving Legal Standards for the Admissibility of Scientific Evidence*, 239 Science 1508, 1512 (1988); *Coppolino v. State*, 223 So.2d 68, 70 (Fla.App. 2 Dist. 1968)("Where the evidence is based solely upon scientific tests and experiments, it is essential that the reliability of the tests and results thereof shall be recognized and accepted by scientists").
[427] *Daubert v. Merrell Dow Pharmaceuticals, Inc.*, 509 U.S. 579, 590 (1993); *Chapman v. Maytag Corp.*, 297 F.3d 682, 688 (7th Cir. 2002)("A very significant *Daubert* factor is whether the proffered scientific theory has been subjected to the scientific method").
[428] *Daubert v. Merrell Dow Pharmaceuticals, Inc.*, 509 U.S. 579, 590 (1993).
[429] *Frye v. United States*, 293 F. 1013 (1923). *Frye* is not covered independently. Instead, *Frye* regimes are addressed somewhat awkwardly under the general acceptability prong of *Daubert* and state based EVID. R. 702 for ease of presentation.
[430] *Daubert v. Merrell Dow Pharmaceuticals, Inc.*, 509 U.S. 579, 590 n.9 (1993).
[431] *Reese v. Stroh*, 874 P.2d 200, 206 (1994).
[432] *State v. Brown*, 687 P.2d 751, 754 (Or. 1984).
[433] *State v. O'Key*, 899 P.2d 663, 679 Fn.24 (Or. 1995); *Daubert v. Merrell Dow Pharmaceuticals, Inc.*, 509 U.S. 579, 593 (1993).
[434] *Daubert v. Merrell Dow Pharmaceuticals, Inc.*, 509 U.S. 579, 593 (1993).
[435] *Daubert v. Merrell Dow Pharmaceuticals, Inc.*, 509 U.S. 579, 592-593 (1993).

accomplishments nor may the courts."[436] "The focus, of course, must be solely on principles and methodology, not on the conclusions that they generate."[437] The following factors are relevant to the determination of scientific reliability.

3. VALIDITY[438]

   a. *Daubert v. Merrell Dow Pharmaceuticals, Inc.*, 509 U.S. 579, 591-592 (1993)(Scientific validity for one purpose is not necessarily scientific validity for other, unrelated purposes).

      i. "'Fit' is not always obvious, and scientific validity for one purpose is not necessarily scientific validity for other, unrelated purposes. The study of the phases of the moon, for example, may provide valid scientific 'knowledge' about whether a certain night was dark, and if darkness is a fact in issue, the knowledge will assist the trier of fact. However (absent creditable grounds supporting such a link), evidence that the moon was full on a certain night will not assist the trier of fact in determining whether an individual was unusually likely to have behaved irrationally on that night. Rule 702's 'helpfulness' standard requires a valid scientific connection to the pertinent inquiry as a precondition to admissibility."

   b. *State v. Lasworth*, 42 P.3d 844, 847-848 (N.M.App. 2001) i. (Scientific validity for one purpose is not necessarily scientific validity for other, unrelated purposes); ii. (Improperly designed validation study does not permit establishment of validity).

      i. "Before scientific evidence may be admitted, the proponent must satisfy the trial court that the technique used to derive the evidence has scientific validity-there must be 'proof of the technique's ability to show what it purports to show'…As Dr. Burns has observed, 'the objective of the test is to discriminate between drivers above and below the statutory BAC limit, *not to measure driving impairment.*' Based on Dr. Burns' testimony and our own review of the 1995 Colorado Report, as well as her published statements, we conclude that the HGN FST has not been scientifically validated as a direct measure of impairment. We conclude that the sole purpose for which the HGN FST arguably has been scientifically validated is to discriminate between drivers above and below the statutory BAC limit."

      ii. "Some minimal level of knowledge of the underlying substantive area of science is necessary even to design a statistical study…The district court appears to have been concerned that without a more detailed understanding of the causes of HGN, the court could not be sure the results obtained by Dr. Burns and other HGN researchers were not a 'coincidence.' We share the district court's concern…At the time of the Colorado study, a BAC of 0.05 percent or greater provided grounds for arrest under Colorado law. The mean BAC of the 234 motorists was 0.152 percent, or *over three times* the statutory limit under Colorado law. Of the 234 motorists, 184 had BACs at or above the statutory limit of 0.05 percent; and, of these 184 motorists, 133 had BACs at or

[436] *U.S. v Horn*, 185 F.Supp.2d 530, 536 Fn.15 (D.Md. 2002).
[437] *Daubert v. Merrell Dow Pharmaceuticals, Inc.*, 509 U.S. 579, 595 (1993). See, however, *General Electric Co. v. Joiner*, 522 U.S. 136, 146 (1997)("A court may conclude that there is simply too great an analytical gap between the data and the opinion proffered.").
[438] *Daubert v. Merrell Dow Pharmaceuticals, Inc.*, 509 U.S. 579, 591 (1993).

above 0.10, or *over twice* the statutory limit. The driving behaviors that led the officers participating in the study to stop a motorist in the first place clearly were selecting out of the general driving population a highly intoxicated group of test subjects. If the officers had simply arrested every one of the 234 motorists, without even administering the FSTs, seventy-nine percent (184 of 234) of their arrest-release decisions would have been correct. In the actual study, the researchers concluded that arrest-release decisions based on the FSTs were correct eighty-six percent of the time. Thus, administration of the FSTs did not dramatically improve the overall percentage of correct decisions. Further, among motorists whose BACs fell in the range between 0.03 to 0.07 percent (0.05 percent ± 0.02 percent), arrest-release decisions based on the FSTs were correct only 57 percent (21 of 37) of the time. We share the district court's concern that some coincidental factor, such as the driving behaviors that led an officer to stop a motorist in the first place, were largely responsible for the claimed ability of the FSTs to discriminate between motorists above and below the statutory BAC."

4. PUBLICATION IN PEER REVIEWED JOURNAL[439]

   a. *People v. Smith*, 132 Cal.Rptr.2d 230, 249 (Cal.App. 2 Dist. 2003)(Result of peer reviewed published NIST interlaboratory studies evidence of general acceptability).

      i. "In an article appearing in the Journal of Forensic Sciences, *NIST Mixed Stain Studies # 1 and # 2: Interlaboratory Comparison of DNA Quantification Practice and Short Tandem Repeat Multiplex Performance with Multiple-Source Samples*, two interlaboratory comparison exercises conducted by the National Institute of Standards and Technology concluded: "Given an appropriate total amount of DNA in the reaction mixture, current STR multiplex systems reliably amplify multiple-source DNA." (Duewer, *NIST Mixed Stain Studies # 1 and # 2: Interlaboratory Comparison of DNA Quantification Practice and Short Tandem Repeat Multiplex Performance With Multiple-Source Samples* (2001) 46 J. Forensic Sci. 1199, 1209.)…Judge Fulgoni's finding that the mixed sample analysis…is accepted by the scientific community was well-reasoned."

5. KNOWN OR POTENTIAL RATE OF ERROR (UNCERTAINTY)[440]

   a. *U.S. v. Allison*, 63 M.J. 365, 369-370 (2006)(Necessity of uncertainty to meaning of result).

      i. "…evidence of statistical probabilities is not only 'basic to DNA analysis,' but also essential to the admissibility of that analysis. In this regard, we follow the state courts which have held that without evidence of statistical frequencies, DNA evidence is meaningless and would not be admissible…The record reflects that Mr. Y and Miss J had received training in DNA statistical analysis and both had considerable experience in conducting that analysis…Both experts responded to questions regarding their statistical conclusions and their understanding of the databases upon which their calculations relied. The testimony also established that the method of calculation

[439] *Daubert v. Merrell Dow Pharmaceuticals, Inc.*, 509 U.S. 579, 594 (1993).
[440] *Daubert v. Merrell Dow Pharmaceuticals, Inc.*, 509 U.S. 579, 594 (1993).

utilized in the analysis had been developed by statisticians and was widely accepted…We therefore conclude that the military judge did not abuse his discretion in allowing the witnesses to testify regarding the statistical frequencies establishing the relevance of the DNA evidence."

b. *DeLuca by DeLuca v. Merrell Dow Pharmaceuticals, Inc.*, 911 F.2d 941, 955-956 (3rd Cir. 1990)(Importance of uncertainty to meaning of result).

   i. "We stress at the outset that the confidence level or 'significance' of a statistical analysis is but a part of a meaningful evaluation of its reliability…any assessment of reliability under Section 702 should be conducted with an eye to all the risks of error posed by the proffered evidence…The root issue it poses is what risk of what type of error the judicial system is willing to tolerate…[courts] may consider the extent to which members of these communities decline to give any weight to inferences not supported by [a particular] statistical significance."

c. *U.S. v. Downing*, 753 F.2d 1224, 1239 (3rd Cir. 1985)(Importance of error rates to reliability).

   i. "The frequency with which a technique leads to erroneous results will be another important component of reliability. At one extreme, a technique that yields correct results less often than it yields erroneous one is so unreliable that it is bound to be unhelpful to a finder of fact. Conversely, a very low rate of error strongly indicates a high degree of reliability. In addition to the rate of error, the court might examine the type of error generated by a technique."

d. *Thomas v. Allen*, 614 F.Supp.2d 1257, 1268-1281 (N.D.Ala. 2009)(Necessity of uncertainty to meaning of result).

   i. "A key task for the...analyst applying a scientific method to conduct a particular analysis, is to identify as many sources of error as possible, to control or to eliminate as many as possible, and to estimate the magnitude of remaining errors so that the conclusions drawn from the study are valid. National Research Council, Strengthening Forensic Science in the United States: A Path Forward, Chap. 4, at 5 (Washington: The National Academies Press 2009)."

   ii. "A critical question that must be addressed is: '*How much confidence can this court place in the IQ scores produced by the tests administered to petitioner?*' Even though most of the intelligence tests that will be discussed later in this opinion are generally considered to be reliable assessment instruments that produce valid IQ scores, there still exists an inherent potential for 'measurement error.' Measurement errors can be either random or systematic. 'Random errors' are caused by any factors that randomly affect measurement of test variables…The important attribute of random errors is that they do not have consistent effects across the entire population of persons to whom the test instrument is administered."

   iii. "'Systematic errors,' on the other hand, are test-specific sources of error that are caused by any factors that *systematically* affect IQ measurements across the entire population

of test subjects. Systematic errors also can be generated by many variables, but usually they can be traced to inadequacies in the assessment instrument itself. Unlike random errors, systematic errors tend to have consistently positive or negative effects upon the performance scores generated by each individual to whom the test is administered. To use a pedestrian example, suppose 'you recorded the temperature every day in your backyard. If your thermometer was incorrectly calibrated, so that it was always 4 degrees too high, the faulty thermometer would produce a systematic error (an upward bias) in your measurement.'"

iv. "A 'true' IQ score is the hypothetical score a test subject would obtain if no measurement error influenced his or her performance during the administration of an intelligence assessment instrument. No clinician, much less this court, can state a test subject's 'true' score with *absolute* certainty, because error *always* is present in any testing situation…Every intelligence test has a [Standard Error of Measurement], which is used to calculate a range of scores lying along a continuum (think of a yardstick), and evenly arranged on each side of the IQ score obtained during an individual administration of the test. The test subject's 'true' IQ most likely lies within that range above and below his or her actual test score."

v. "The attorneys for both parties and their expert witnesses stipulated that a standard error of measurement in the neighborhood of approximately ± 5 points is proper for full-scale IQ test scores produced by the intelligence assessment instruments discussed in this opinion. The American Psychiatric Association agrees: the most recent edition of its *Diagnostic and Statistical Manual of Mental Disorders* notes that 'there is a measurement error of approximately 5 points in assessing IQ'…even though the legal cut-off score for a finding of 'significantly subaverage intellectual functioning' is stated in opinions of the Alabama Supreme Court as 'an IQ of 70 or below,' a court should not look at a raw IQ score as a *precise* measurement of intellectual functioning. A court must also consider…the standard error of measurement in determining whether a petitioner's IQ score falls within a *range* containing scores that are less than 70."

e. *Henricksen v. ConocoPhillips Co.*, __ F.Supp.2d __, (E.D.Wash. 2009)(Poor methodology in determining error rate undermines reliability).

i. "The court also considers the potential rate of error. *Nordlinder* was not an appropriately designed study to yield reliable or conclusive results on the difference between benzene exposures in open and closed terminals. The small sample sizes of five and sixteen leaves a great deal of uncertainty about the measurements obtained. If in error, Henricksen's cumulative dose calculation could be off by 500%. Kaltofen's methodology in arriving at the multiplier of 5 shows a lack of scientific rigor in that he expands the application of *Nordlinder* beyond good science, drawing conclusions the authors of the study did not make from limited data. It is this kind of scientifically unsupported 'leap of faith' which is condemned by Daubert."

f. *Phillips v. Raymond Corp.*, 364 F.Supp.2d 730, 741 (N.D.Ill. 2005)(Inability to document error rate undermines conclusion of reliability).

i. "The Court notes that…it appears that the potential rate of error of Liu's calculations is unknown. Apparently, for Liu to be able to determine the rate of error for his tests (thus helping to make them scientifically valid), he would have had to engage in a 'retrospective analysis.' Liu did not conduct such a 'retrospective analysis.' Thus, Liu cannot provide a potential rate of error. This cuts against admissibility."

g. E.E.O.C. v. Ethan Allen, Inc., 259 F.Supp.2d 625, 634-636 (N.D.Ohio 2003)(Method that can only yield only a 68% level of confidence in its conclusion is not reliable).[441]

i. "That a small number of analysts got together and agreed that statistical significance in ink dating is acceptable at the level of one standard deviation, however, does not make it so. It is an elementary statistical truth that a test using one standard deviation ('1STD') as its measure of statistical significance yields a 68% confidence level in the results. This is the same as saying there is about a one in three chance that the test results are not significant at all…In comparison, a test that uses a 2STD measure of statistical significance yields a 95% confidence level in the results, and a test that uses a 3STD measure of statistical significance yields a 99.7% confidence level in the results. By adopting the 1STD measure, Speckin and his SOFIA cohorts agreed that their ink-dating tests would be only moderately sensitive to error, even assuming completely 'logical' data…Because Speckin has used a 1STD measure of statistical significance, he simply cannot validly opine '*to a high degree of scientific certainty*' that the Mora letter was written within the last 3 1/2 years, and not in 1994. A high degree of scientific certainty *may* be attained by tests using a 2STD measure of statistical significance, but the confidence level underlying Speckin's results is only slightly higher than the predicted results of tossing a coin. Unsurprisingly, the *Wang* court concluded that its own confidence level in Speckin's opinion could only be 'weak,' and noted that using a test with a sensitivity of only 1STD is a 'departure from the accepted norms of analytical chemistry'…In sum, the statistical analysis used by Speckin to reach his expert opinion does not support that very opinion. *Daubert* instructs that this court 'should consider the known or potential rate of error" in the tests underlying an expert's conclusion. The rate of error of Speckin's analysis, given his use of a 1STD measure of statistical significance, is about one out of three. As such, the Court finds that any expert testimony offered by Speckin regarding ink-dating using relative ink age comparison tests cannot properly be admitted as an expert conclusion."

h. U.S. v. Shea, 957 F.Supp. 331, 341-343 (D.N.H. 1997)(recognizing disparate estimates of uncertainty may exist within scientific community).

i. "The government's estimate of a 1 in 200,000 random match probability is based primarily on information drawn from a PCR database comprised of DNA profiles for 148 Caucasians, 145 African Americans, 94 Southeastern Hispanics, and 96 Southwestern Hispanics. Shea contends that this database is simply too small to be used reliably in estimating random match probabilities with the product rule…legitimate questions can be raised concerning the reliability of a random match probability that is

[441] The focus is on the final result here. The author did not have access to the underlying evidence or testimony and suspects that there may be a misunderstanding in the court's statistical analysis. The discussion is useful whether strictly technically correct or not though.

estimated with the product rule from a database as small as the one used here. Because such databases are comprised of a limited number of samples, the possibility of random error ordinarily must be considered. Further, legitimate questions can be raised concerning the power of existing statistical methods to detect deviations from Hardy-Weinberg and linkage equilibrium when small databases are used. If random error is not accounted for and if the likely potential effects of factors such as population substructuring are not identified and addressed, a random match probability estimated with the product rule may be unreliable…Undetected population substructuring and random error can also affect individual random match probability calculations in ways that are difficult to predict…Whether the adjustments to the product rule suggested in the NRC II report are sufficiently conservative and whether a database of 148 is of sufficient size to serve as the basis for a reliable random match probability estimate are important questions about which population geneticists can legitimately disagree. However, Rule 702 does not require scientific consensus. The government has produced a peer-reviewed study using accepted statistical methods to support its position that the estimation of a random match probability from the database used in this case will produce a reliable result. It has further qualified its estimate in accordance with the recommendations of a distinguished committee of scientists and academicians that included leading population geneticists as members. Under these circumstances, the concerns raised by Dr. Shields affect the weight that should be given to the evidence rather than its admissibility."

i. *State v. Morales*, 45 P.3d 406, 412 (N.M.App. 2002)(Importance of error rate and/or sources of error to determination of reliability).

   i. "Evidentiary reliability has been described as 'the hallmark for the admissibility of scientific knowledge'…In his testimony, Deputy Gonzales acknowledged that he knew nothing about the chemical features of the field test and how it produced a certain color that identified heroin. The deputy also had no scientific evidence about the percentage reliability of the field test. Instead, the State relies exclusively on the deputy's own testimony that the field test was reliable. Clearly, this will not do. Our Supreme Court pointed out…that 'if police officers are not qualified to testify about the scientific bases underlying the...test, they are not competent to establish that the test satisfies the relevant admissibility standard'…the State has the burden to establish the validity of the scientific principles on which the test is based and its scientific reliability when the State elects to rely on a field test to prove the identity of the contraband. We further hold that testimony by a law enforcement officer will not, without more, be sufficient to support admission of the results, when the officer cannot explain the scientific principles that the test uses, the percentage of false positives or negatives that the test will produce, or the factors that may produce those false results."

j. *Ramirez v. State*, 810 So.2d 836, 849-851 (Fla. 2001)(Claim of infallibility, i.e. an error rate of zero, undermines conclusion of general acceptability).

   i. "…Hart's testing procedure possesses none of the hallmarks of acceptability that apply in the relevant scientific community to this type of evidence. This is particularly true in light of the extraordinarily precise claims of identification that Hart makes under his

testing procedure-i.e., he claims that a 'match' made pursuant to his method is made with absolute certainty. Such certainty, which exceeds even that of DNA testing, warrants careful scrutiny in a criminal-indeed, a capital-proceeding. First, the record does not show that Hart's methodology-and particularly his claim of infallibility-has ever been formally tested or otherwise verified…Fourth, the record does not show that the error rate for Hart's method has ever been quantified. On the contrary, the State's experts testified that the method is infallible, that it is impossible to make a false positive identification."

k. *State v. Cauthron*, 846 P.2d 502, 906-907 (Wash. 1993)(Failure to provide probability of results rendered evidence not generally accepted under *Frye* or helpful to finder of fact under ER 702).

   i. "The expert testimony here did not provide any probability statistics. Instead, four experts testified that Cauthron's DNA 'matched' the semen samples taken from the victims…This testimony should not have been admitted, because it does not meet the test for expert testimony. As stated above, expert testimony is admissible only when the underlying scientific principle satisfies the threshold *Frye* requirements *and* the testimony meets the 2-part test of ER 702:…and (2) the expert testimony would be helpful to the finder of fact…Because the testimony presented did not include the background probability information, it was insufficient…Testimony of a match in DNA samples, without the statistical background or probability estimates, is neither based on a generally accepted scientific theory nor helpful to the trier of fact."

l. *Nelson v. State*, 628 A.2d 69, 76 (Del. 1993)(To say that two DNA patterns match, without providing any scientifically valid estimate of the frequency with which such matches might occur by chance, is meaningless).

   i. "We find the Superior Court's rationale for admitting the State's evidence of a match while excluding its proffered statistical interpretation of the match to be flawed. The court excluded such evidence not on its own merits or for a found lack of reliability, but out of concern that the statistics would be overly prejudicial to the defendant and possibly confusing or misleading to the jury. The court's reference to Nelson's indigency seems misplaced, in the absence of any record evidence of an application for funds to employ an expert. In any event, we find the court's ruling inherently inconsistent since, without the necessary statistical calculations, the evidence of the match was "meaningless" to the jury and, thus, inadmissible."

m. *State v. Brown*, 470 N.W.2d 30, 33 (Iowa, 1991)("Without statistical evidence, the ultimate results of DNA testing would become a matter of speculation").

   i. "Brown contends that statistical probabilities could have been determined by the jury without the assistance of an expert. However, the test for admission of expert testimony is not whether the jury might be able to arrive at the same conclusion but whether the evidence in question will assist the jury. *See* Iowa R.Evid. 702. In the present case, it is doubtful that jurors could take the probabilities of the four separate segments, combine them, and arrive at an answer with any degree of certainty as to its correctness.

Furnishing statistical analysis would assist the trier of fact in such a case and that is the heart of admissibility under rule 702. Without statistical evidence, the ultimate results of DNA testing would become a matter of speculation."

n. Com. v. Curnin, 565 N.E.2d 440, 442-443 n.7 (Mass. 1991)(Failure to provide rational basis for probability of results rendered evidence not generally accepted under *Frye*).

   i. "[T]here is no demonstrated general acceptance or inherent rationality of the process by which Cellmark arrived at its conclusion that one Caucasian in 59,000,000 would have the DNA components disclosed by the test that showed an identity between the defendant's DNA and that found on the nightgown…we would not permit the admission of test results showing a DNA match (a positive result) without telling the jury anything about the likelihood of that match occurring…The evidence and other material that may appropriately be considered do not warrant the conclusion that Cellmark followed a generally accepted or obviously logical procedure in deciding the likelihood that someone else would have the same DNA characteristics as those that were identified in the comparison test."

o. State v. Keller, 36 Wn.App. 110, 113-114 (1983)("[T]he margin of error in the Breathalyzer should be considered by the trier of fact in deciding whether the evidence sustains a finding of guilt beyond a reasonable doubt").

   i. "…a Breathalyzer reading of .10 percent is not conclusive proof of guilt. The State still has the burden of proving beyond a reasonable doubt that the .10 reading is correct, and the defendant may attack the accuracy of the reading…The foregoing suggests that the margin of error in the Breathalyzer should be considered by the trier of fact in deciding whether the evidence sustains a finding of guilt beyond a reasonable doubt. The weight to be given the Breathalyzer reading is left to the trier of fact, as is the weight to be accorded other evidence in the case. The trial court considered all the evidence, including the Breathalyzer's margin of error, and made a factual determination that Keller's violation of the statute was established beyond a reasonable doubt."

p. State v. Boehmer, 613 P.2d 916, 918-919 (Haw. App. 1980)(State cannot prove that BAC is greater than the legal limit without accounting for the margin of error).

   i. "In both of the cases at bar, the State has failed to establish a critical fact. The State merely demonstrated that the reading of the breathalyzer machine was 0.10% for Defendant Boehmer and 0.11% for Defendant Gogo. The inherent margin of error could put both defendants' actual blood alcohol level below the level necessary for the presumption to arise. The failure of the prosecution to establish beyond a reasonable doubt that the actual weight of alcohol in defendants' blood was at least .10% required the trial judge to ignore [any presumption based on the test result]."

q. State v. Bjornsen, 271 N.W.2d 839, 840 (Neb. 1978)(State cannot prove that BAC is greater than the legal limit without accounting for the margin of error).

i. "The Legislature has selected a particular percent of alcohol to be a criminal offense if present in a person operating a motor vehicle. It is not unreasonable to require that the test, designed to show that percent, do so outside of any error or tolerance inherent in the testing process."

6. STANDARDS CONTROLLING THE TECHNIQUE'S OPERATION[442]

a. *U.S. v. Prime*, 431 F.3d 1147, 1153-1154 (9th Cir. 2005)(ASCLD accreditation and utilization of methods established by ASTM evidence of reliability).

i. "The court recognized that although this area has not been completely standardized, it is moving in the right direction. The Secret Service laboratory where Storer works has maintained its accreditation with the American Society of Crime Laboratory Directors since 1998, based on an external proficiency test. Furthermore, the standard nine-point scale used to express the degree to which the examiner believes the handwriting samples match was established under the auspices of the American Society for Testing and Materials ('ASTM'). The court reasonably concluded that any lack of standardization is not in and of itself a bar to admissibility in court."

b. *Alfred v. Caterpillar, Inc.*, 262 F.3d 1083, 1087-1088 (10th Cir. 2001)(Testimony based on SAE engineering standards evidence of reliability and departure from standards relevant).

i. "Munsell's testimony was based on engineering standards promulgated by the Society of Automotive Engineers ("SAE") as well as on his investigative work. Citing SAE Standard J297, entitled "Operator Controls on Industrial Equipment," he opined that the variable speed control on a paver should be in the form of a lever rather than a rotary dial. Because the paver involved in the litigation was equipped with a rotary dial instead of a lever, he concluded, its design was defective for failing to meet the SAE standard. Munsell testified that he had nine years of experience…and that he has routinely researched and applied engineering standards promulgated by various organizations, including the SAE. He testified further regarding his methodology in this case, which included researching engineering standards…and applying those standards to knowledge gained during field research. Defendant did not dispute that the SAE standards upon which Munsell's opinion was based are well-accepted in the engineering community. Technical committees of the SAE draft and review engineering safety standards for mobility systems, including off-highway equipment. According to Munsell's testimony, several Caterpillar employees were members of committees responsible for promulgating the SAE standards…we are persuaded that Munsell's testimony that the speed control mechanism did not comply with SAE J297 was both reliable and relevant to the issue of defective design. Munsell's testimony was reliable-meeting one of the *Daubert* criteria-because it was the result of his having researched and applied standards promulgated by an internationally recognized organization of engineers. The testimony was relevant-meeting the other-because although it is not dispositive and might be countered by conflicting testimony, it could allow the jury to infer Caterpillar's paver was defective for failing to meet industry

[442] *Daubert v. Merrell Dow Pharmaceuticals, Inc.*, 509 U.S. 579, 594 (1993).

design standards. Because that portion of his opinion qualified as admissible expert testimony under Rule 702 and *Daubert,* we hold that striking Munsell's testimony as to the paver's failure to comply with SAE J297 was an abuse of the trial court's discretion."

c. *Bourelle v. Crown Equipment Corp.*, 220 F.3d 532, 537-538 (7th Cir. 2000)(Departure from ANSI standards included as evidence of unreliability).

   i. "…the appellants ignore the fact that Pacheco never…submitted his alternative design theories to the American National Standards Institute (ANSI), despite the fact that he was aware of the organization."

d. *Bowers v. Norfolk Southern Corp.*, 537 F.Supp.2d 1343, 1374 (M.D.Ga. 2007)(Testimony reliable because based on ISO standards).

   i. "Plaintiff instead contends that Larson's opinions are not based on a reliable foundation…Purporting to employ methods outlined by the ISO (International Organization for Standardization), Larson measured the level of vibration in various locations inside of the locomotive, including on the bottom of the conductor's seat and on the floor directly underneath the conductor's seat. However, Larson did not measure vibration at the seat-back. Plaintiff claims that Larson's failure to measure vibration at the seat-back renders his opinions unreliable and, therefore, subject to exclusion under Rule 702 and *Daubert.*..Notably, Plaintiff does not challenge Larson's use of the ISO standards. Instead, Plaintiff alleges that, by failing to measure vibration at the seat-back, Larson failed to properly apply those standards. The ISO has promulgated standards for measuring vibration forces on the human body…The ISO procedures for measuring vibration vary according to the position of the person on which the vibration forces are acting and the purpose for which the measurements are taken. For instance, for seated persons, the ISO standards recommend measuring vibration in the following three areas: "the supporting seat surface, the seat-back, and the feet." ISO Standard 2631-1, *Mechanical Vibration and Shock: Evaluation of Human Exposure to Whole-Body Vibration* § 5.3 (1997). For persons in a recumbent position, meanwhile, measurements are taken in different areas, namely, under the pelvis, back, and head. *Id.* Larson concedes that he measured vibration forces at only two of the three recommended areas. He argues, however, that measurement at the seat-back, though recommended by the ISO, was unnecessary, because the ISO standards do not require such measurement for purposes of assessing the effect of vibration on human health…Larson's explanation is supported by the ISO standards. The clause describing the methods for evaluating the effect of vibration on health states: 'measurements...on the backrest...are encouraged. *However, considering the shortage of evidence showing the effect of this motion on health, it is not included in the assessment of the vibration severity.'* ISO Standard 2631-1, *Mechanical Vibration and Shock: Evaluation of Human Exposure to Whole-Body Vibration* § 7.2.3 (1997) (emphasis added). Thus, according to the ISO standards, a seat-back measurement is neither necessary nor helpful…because Larson properly applied internationally-recognized standards, adhering to the guidelines articulated within those standards, his opinions are reliable under *Daubert* and Rule 702."

e. *Milanowicz v. The Raymond Corp.*, 148 F.Supp.2d 525, 533 (D.N.J. 2001)(Referenced to standards published by independent standards organizations such as ASME and ASTM evidence of reliability).

   i. "*Rule 702,* Daubert, *and* Kumho…Independent Standards Organizations-Courts should also examine whether the expert has referenced standards published by independent standards organizations such as the American National Standards Institute (ANSI), Underwriters' Laboratories (UL), the American Society of Mechanical Engineers (ASME), and the American Society for Testing and Materials (ASTM). While lacking the legal authority of federal regulations, they provide detailed design standards which reflect systematic testing and safety certification."

f. *Ex parte Taylor*, 825 So.2d 769, 778 (Ala. 2002)(Use of NIST reference material to validate test method evidence of reliability).

   i. "The DNA analyst's testimony of the NIST sample validations and the positive and negative controls performed on the Perkin-Elmer kits tended to prove their scientific reliability. That is, each NIST sample validation and each positive control demonstrated that the kits could accurately identify a DNA sample of known identity, and each negative control demonstrated that the kits would not indicate identifiable DNA in the absence of DNA…The combination of (1) his explanations of the NIST sample validations and the positive and negative controls, (2) his testimony to the *Daubert/Turner* reliability factors, and (3) his general explanation of the operation of the Perkin-Elmer kits, sufficed to carry the burden of the State to prove the scientific reliability of the kits."

g. *People v. Shreck*, 22 P.3d 68, (Colo. 2001)(Determination by NIST in favor of method evidence of reliability).

   i. "Similarly…the National Institute of Standards and Technology ('NIST') has determined that there are several advantages of using STRs over conventional techniques, and that the use of STRs for genetic mapping and identity testing has become widespread among DNA typing laboratories. John M. Butler & Dennis J. Reeder, *Short Tandem Repeat DNA Internet Database*, http://www.cstl.nist.gov/biotech/strbase/intro.htm... We are therefore convinced that DNA evidence derived from PCR-based testing, and specifically such evidence derived from the STR method is sufficiently reliable under CRE 702."

h. *Com. v. Wilkins*, 605 A.2d 363, 368 (Pa.Super. 1992)(Absent contrary evidence, adherence to NIST standard established proper use).

   i. "Lastly, Ms. Wilkins maintains that no evidence was introduced that the Speedchek device used to time her speed was properly installed pursuant to PennDot regulations…the Commonwealth introduced into evidence a document from the Commonwealth Department of General Services entitled 'Report of Test for Linear Measures' which certified that the linear measurement for the Speedchek tapes conforms to the specifications of the Linear Measure Code of the National Institute of

Standards and Technology (NIST) 'and are correct for law enforcement applications'…The linear measurement which the Department of General Services certified as being in conformity with the NIST standards is five feet. Significantly, counsel for Ms. Wilkins stipulated to the introduction of this Exhibit into evidence and never questioned the correctness of the information contained therein…Therefore, we are unable to conclude that relief is warranted here."

7. GENERAL ACCEPTANCE[443]

a. *Srail v. Village of Lisle*, 249 F.R.D. 544, 562 (N.D.Ill. 2008)(Standards in NFPA handbook evidence methodology generally accepted in relevant scientific community).

i. "The Court finds that Gasser's methodology is sufficiently reliable for the Court to consider his report…Though Lisle challenges the adequacy of this sampling and the use of a random selection process, the NFPA publication on fire flow testing makes no particular recommendations as to what percentage of hydrants in a given area should be tested or how those hydrants should be selected. Rather, it states that 'a group of test hydrants in the vicinity is selected' and that the 'number of hydrants to be used in any test depends upon the strength of the distribution system in the vicinity of the test location.' NFPA Recommended Practice for Fire Flow Testing and Marking of Hydrants, Chapter 4 (2007), 4.3.1 & 4.3.5. Thus the manner of Gasser's selection of hydrants does not suggest that the selection process was flawed or that it failed to meet recognized standards…the fact that Gasser's test procedure was consistent with general industry standards and practices as described in the NFPA handbook supports the proposition that the methodology Gasser employed enjoys general acceptance in the relevant scientific community."

b. *Phillips v. Raymond Corp.*, 364 F.Supp.2d 730, 741 (N.D.Ill. 2005)(ISO and SAE standards are helpful in determining general acceptability of scientific methodology).

i. "Also unhelpful to Phillips is the issue of whether Liu meets the fourth *Daubert* factor-general acceptance. Phillips asserts that Liu meets this prong of the *Daubert* review. (arguing that Liu's testing meets all the standards of the Society of Automotive Engineers and the International Organization for Standardization)…Liu provides nothing more than his own opinion as to the acceptability of his own tests. It would have been helpful if Phillips had demonstrated what the SAE or ISO standards are, for example. Unsubstantiated testimony, such as this, does not ensure that 'the expert's opinion has a reliable basis in knowledge and experience of his discipline.'"

c. *Coffey v. Dowley Mfg., Inc.*, 187 F.Supp.2d 958, 978 (M.D.Tenn. 2002)(Failure to comply with ASTM standards constitutes evidence methods not generally accepted).

i. "Second, the Supreme Court opined that the 'general acceptance' of a theory can have a bearing on the court's Rule 702 inquiry…Dr. Wilson failed to comply with various American Society for Testing and Materials (ASTM) standards [ASTM E 1188-95,

[443] *Daubert v. Merrell Dow Pharmaceuticals, Inc.*, 509 U.S. 579, 594 (1993).

Standard Practice for Collection and Preservation of Information and Physical Items by a Technical Advisor, ASTM E 860-97, *Standard Practice for Examining and Testing Items that are or may become Involved in Litigation*), and ASTM E 678-98, *Standard Practice for Evaluation of Technical Data*]. Dr. Wilson is a member of ASTM, and recognized the authoritative nature of the ASTM standards. His failure to comply with ASTM standards belies Dr. Wilson's claim that his theories are generally accepted."

d. *City of Seattle v. Clark-Munoz*, 93 P.3d 141 (Wash. 2004)(NIST supplies generally accepted definition of traceability).

i. "The question before us is whether these machines have been 'properly checked.' This hinges on the meaning of the term 'traceable.' If 'traceable' is given the scientific meaning articulated by NIST, which requires that uncertainties be noted at each level of removal so that the ultimate uncertainty is known, then the testing machines have not been properly checked. If traceable is given a nonscientific meaning, they may comply…The state toxicologist did not define 'traceable' in the regulations. The NIST policy on traceability outlines the procedures required for traceability…We will give weight to the technical definition of a technical term promulgated by an expert agency…In addition to having a policy on 'traceability,' NIST: 'Adopts for its own use and recommends for use by others the definition of traceability provided in the most recent version of the *International Vocabulary of Basic and General Terms in Metrology:* '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.*' NIST POLICY ON TRACEABILITY, available at http:// ts.nist.gov/traceability/nist % 20traceability% 20 policy-external.htm (quoting INTERNATIONAL VOCABULARY OF BASIC AND GENERAL TERMS IN METROLOGY (VIM), Definition 6.10, BIPM, IEC, IFCC, ISO, IUPAC, IUPAP, OIML, (2d ed., 1993)). This is substantially the definition given by Dr. Ashley Emery, Ph.D, a University of Washington professor and expert witness in the science of metrology (the study of measurements). He testified that the term 'traceable' in science had 'an internationally agreed upon scientific meaning' that included a requirement that the uncertainties at each step be measured. He testified that the requirement that uncertainties be measured and recorded is a critical element of the NIST definition. Further, Dr. Emery testified that '[w]ithout a statement of uncertainty, the measurement is worthless,' and that every scientist would define 'traceable' in these technical terms. The state toxicologist was unaware of the NIST's technical scientific definition when the regulation before us was promulgated. He testified that he did not intend to incorporate it into the breath test regulations. However, while the state toxicologist may not have known the precise definition, he did know it was a term of art: 'The concept of traceability to a reference standard is a common principle in measurement science. It describes the notion that there is an absolute standard for temperature, maintained by the National Institute for Standards and [Technology] (NIST), and that the reference thermometer used to certify the mercury in glass thermometers used in this program, must be compared against a thermometer which has been checked either directly or indirectly against that absolute standard, and thus can be 'traced' to it'…All this weighs in favor of our conclusion that 'traceable' is a technical term, to be given its technical meaning. As Judges Chapman,

Eiler, and Jacke found in their well reasoned opinion: 'If the citizens of the State of Washington are to have any confidence in the breath testing program, that program has to have some credence in the scientific community as a whole.'"

e. *Lemour v. State*, 802 So.2d 402, 406 (Fla.App. 2001)(Official NIST statement evidence of general acceptability).

i. "Furthermore, the National Institute of Standards and Technology [NIST] website reflects that 'multiplex STRs are used extensively in the forensic field, [and] NIST has concluded that 'multiplex [testing]...is an ideal technique for DNA typing....'"

f. *State v. Copeland*, 922 P.2d 1304, 1316 (1996)(Official report by National Academy of Sciences authoritative in setting forth proper scientific practices).

i. "FN1. The scientific explanation here is drawn primarily from Committee on DNA Technology in Forensic Science, *DNA Technology in Forensic Science* (National Academy Press 1992) (*DNA Technology)*…"

ii. "The court in *Cauthron* relied considerably upon conclusions drawn by a "committee of eminent scientists and jurists" (the Committee) which had researched and analyzed the status of forensic DNA typing under the auspices of the National Academy of Sciences. Committee on DNA Technology in Forensic Science, *DNA Technology in Forensic Science* (National Academy Press 1992) (*DNA Technology*)."

g. *State v. Cauthron*, 846 P.2d 502, 504 (Wash. 1993)(Official report by National Academy of Sciences authoritative in setting forth proper scientific practices).

i. Decision relied upon report by the National Academy of Sciences Committee on DNA Technology in Forensic Science, "[a] committee of eminent scientists and jurists [who have] exhaustively researched and analyzed the current status of forensic DNA typing."

C. STATUTORY/REGULATORY METROLOGICAL PROVISIONS

1. FEDERAL – NATIONAL INSTITUTE OF STANDARDS AND TECHNOLOGY (NIST)

a. 15 USCA § 271 – Findings, declarations and purpose.

i. "The Congress finds and declares the following…Precise measurements, calibrations, and standards help United States industry and manufacturing concerns compete strongly in world markets. Improvements in manufacturing and product technology depend on fundamental scientific and engineering research to develop the precise and accurate measurement methods and measurement standards needed to improve quality and reliability…Scientific progress, public safety, and product compatibility and standardization also depend on the development of precise measurement methods, standards, and related basic technologies…The Federal Government should maintain a national science, engineering, and technology laboratory which provides measurement methods, standards, and associated technologies…Such national laboratory also should serve industry, trade associations, State technology programs, labor organizations,

professional societies, and educational institutions by disseminating information on new basic technologies…"

ii. "It is the purpose of this chapter to rename the National Bureau of Standards as the National Institute of Standards and Technology and to modernize and restructure that agency…while maintaining its traditional function as lead national laboratory for providing the measurements, calibrations, and quality assurance techniques which underpin United States commerce, technological progress, improved product reliability and manufacturing processes, and public safety."

b. 15 USCA § 272 – Functions and activities.

i. Functions: "…to develop, maintain, and retain custody of the national standards of measurement, and provide the means and methods for making measurements consistent with those standards; to compare standards used in scientific investigations, engineering, manufacturing, commerce, industry, and educational institutions…to provide United States industry, Government, and educational institutions with a national clearinghouse of current information, techniques, and advice…to assist industry in the development of measurements, measurement methods, and basic measurement technology; to determine, compile, evaluate, and disseminate physical constants and the properties and performance of conventional and advanced materials when they are important to science, engineering, manufacturing, education, commerce, and industry…to develop a fundamental basis and methods for testing materials, mechanisms, structures, equipment, and systems…to cooperate…in establishing standard practices, codes, specifications, and voluntary consensus standards…to coordinate Federal, State, and local technical standards activities and conformity assessment activities, with private sector technical standards activities and conformity assessment activities."

ii. Activities: "…construct physical standards; test, calibrate, and certify standards and standard measuring apparatus; study and improve instruments, measurement methods, and industrial process control and quality assurance techniques; cooperate with the States in securing uniformity in weights and measures laws and methods of inspection; cooperate with foreign scientific and technical institutions to understand technological developments in other countries better; prepare, certify, and sell standard reference materials for use in ensuring the accuracy of chemical analyses and measurements of physical and other properties of materials… undertake such research in engineering, pure and applied mathematics, statistics, computer science, materials science, and the physical sciences as may be necessary to carry out and support the functions specified in this section; compile, evaluate, publish, and otherwise disseminate general, specific and technical data resulting from the performance of the functions specified in this section or from other sources when such data are important to science, engineering, or industry, or to the general public, and are not available elsewhere; collect, create, analyze, and maintain specimens of scientific value…evaluate promising inventions and other novel technical concepts."

c. NIST OFFICE OF LAW ENFORCEMENT STANDARDS

i. "[T]he Office of Law Enforcement Standards (OLES) addresses the technology and metrology needs of the criminal justice, public safety, public security and greater homeland security communities. Since 1971, OLES's customers have been corrections personnel, forensic scientists and police officers, firefighters, and others responsible for the safety and security of people and property. Through our work on performance standards for critical technologies such as ballistic body armor, metal detectors, chemical systems and protective equipment, computer forensics, DNA analysis…OLES has developed unique expertise…In addition to developing minimum performance standards, OLES develops reference materials (RMs) and standard reference materials (SRMs) for use in test procedures and to calibrate equipment. OLES develops technology and metrology to support the advancement of equipment and methods used to address the needs of criminal justice, public safety, emergency responder and homeland security agencies. OLES authors equipment user guides; designs methods for examining evidentiary materials; develops technology where appropriate and applicable; and provides technical advice and assistance to agencies throughout the criminal justice, public safety, emergency responder and homeland security communities."[444]

2. STATE – RECOGNITION OF NIST/ISO STANDARDS

a. Most states have their own statutory or regulatory provisions governing weights, measures and standards. For links to a majority of these statutes see http://ts.nist.gov/WeightsAndMeasures/WMLAW.cfm.

b. 15 USCA § 272:

i. NIST "shall work directly with States, local governments, and other appropriate organizations to provide for extended distribution of Standard Reference Materials, Standard Reference Data, calibrations, and related technical services and to help transfer other expertise and technology to the States."

c. STATE METROLOGICAL LABORATORIES:

i. "State legal metrology laboratories are custodians at the State level of measurement standards that serve as the basis for ensuring equity in the marketplace and as reference standards for calibration services for indigenous industry."[445]

a) NIST "has developed performance standards and formalized procedures for Recognition of State legal metrology laboratories on a voluntary basis. Certificates of Measurement Traceability are issued upon evaluation of the laboratory's ability to make reliable metrological measurements."[446]

b) "The general requirements in sections 4 and 5 incorporate ISO/IEC 17025:2005 (as adopted by the NVLAP Calibration Laboratories Accreditation Program) and

[444] NIST, OLES Mission Statement http://www.eeel.nist.gov/oles/oles_mission.html.
[445] NIST, *State Weights and Measures Laboratories, Program Handbook*, NIST HB 143, 11 (2007).
[446] NIST, *State Weights and Measures Laboratories, Program Handbook*, NIST HB 143, 11 (2007).

address internationally accepted quality management practices for calibration and testing laboratories."[447]

ii. NIST CERTIFIED/ACCREDITED STATE WEIGHTS AND MEASURES PROGRAMS: See http://ts.nist.gov/WeightsAndMeasures/statelabcontact.cfm

United States Department of Commerce
National Institute of Standards and Technology

~ Certificate of Measurement Traceability ~
for

Alaska

a participant in the NIST Weights and Measures Measurement Assurance Program for the Year(s)

2009 - 2010

The State standards are traceable to NIST and the laboratory has demonstrated* measurement proficiency through training and interlaboratory comparisons compliant to NIST HB 143:2007 (ISO/IEC 17025:2005).

*Demonstrated Areas of Proficiency

Mass
Echelon III
1000 lb - 0.001 lb
5 kg - 1 mg
8 oz - 0.03125 oz
Weight Carts
5000 lb - 2200 lb
Wheel Load Weighers
20 000 lb - 5 000 lb

Volume
Volume transfer, II
1500 gal - 5 gal

Frequency
Tuning Forks
100 mph - 35 mph

Carol T. Hockert
for the National Institute of Standards and Technology

Chief, Weights and Measures Division
Title

January 1, 2009 to December 31, 2010
Effective Date

3. RECOGNITION OF STANDARDS AS EVIDENCE THAT ADHERENCE IS NECESSARY FOR RELIABILITY

a. Where the government develops, adopts or generally relies upon standards to establish the reliability of measurement procedures and results within its jurisdiction, it should be viewed as evidence that adherence to such standards is necessary to establish the reliability of all such measurement procedures and results.

b. *United States v. Van Griffin*, 874 F.2d 634, 638 (9th Cir. 1989)(Standards issued by agency responsible for subject matter is evidence standards are necessary for reliability).

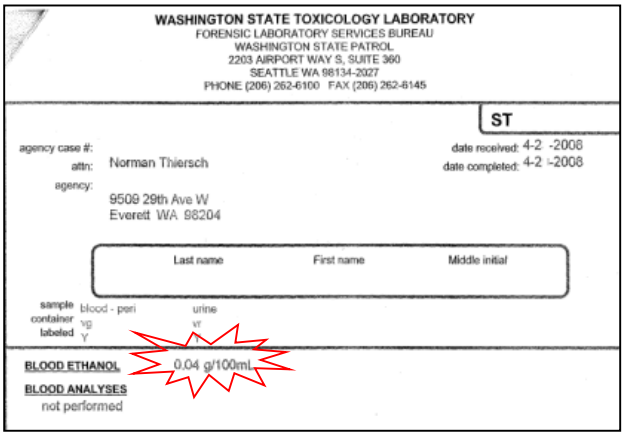
i. "*Admissibility of the Department of Transportation Manual.* The basis on which counsel for Griffin sought to introduce the manual was to impeach Griffin but Ranger Oltrogge testified that he had not relied upon or even ever heard of the manual. The manual therefore was not a challenge to the ranger's testimony and therefore not proper impeachment…The manual, however, could have been introduced by the defendant as part of his defense in order to show the measures that are necessary to be taken in order to have a reliable test for nystagmus. We do not say that every publication of every branch of government of the United States can be treated as a party admission by the United States…In this case the government department charged with the development of rules for highway safety was the relevant and competent section of the government; its pamphlet on sobriety testing was an admissible party admission."

VIII. APPLICATIONS

[447] NIST, *State Weights and Measures Laboratories, Program Handbook*, NIST HB 143, 11 (2007).

A. EVID. R. 702. – BLOOD TESTING: REPORTING RESULTS

1. FAILURE TO REPORT UNCERTAINTY WITH BLOOD TEST RESULTS



WASHINGTON STATE TOXICOLOGY LABORATORY
FORENSIC LABORATORY SERVICES BUREAU
WASHINGTON STATE PATROL
2203 AIRPORT WAY S, SUITE 360
SEATTLE WA 98134-2027
PHONE (206) 262-6100 FAX (206) 262-6145

ST

agency case #:
attn: Norman Thiersch
agency:
9509 29th Ave W
Everett WA 98204

date received: 4-2 -2008
date completed: 4-2 -2008

Last name | First name | Middle initial

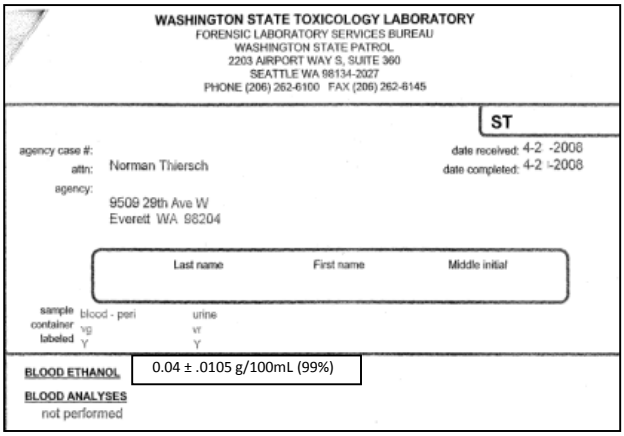
sample container labeled: blood - peri: yg Y; urine: yr Y

BLOOD ETHANOL 0.04 g/100mL

BLOOD ANALYSES
not performed

2. Correct Alternatives

a. COVERAGE INTERVAL



WASHINGTON STATE TOXICOLOGY LABORATORY
FORENSIC LABORATORY SERVICES BUREAU
WASHINGTON STATE PATROL
2203 AIRPORT WAY S, SUITE 360
SEATTLE WA 98134-2027
PHONE (206) 262-6100 FAX (206) 262-6145

ST

agency case #:
attn: Norman Thiersch
agency:
9509 29th Ave W
Everett WA 98204

date received: 4-2 -2008
date completed: 4-2 -2008

Last name | First name | Middle initial

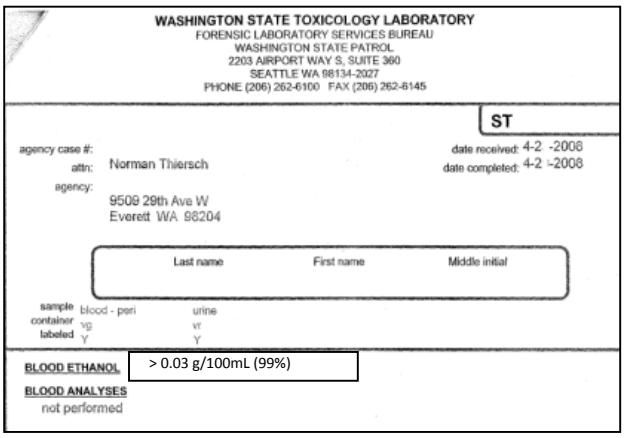
sample container labeled: blood - peri: yg Y; urine: yr Y

BLOOD ETHANOL 0.04 ± .0105 g/100mL (99%)

BLOOD ANALYSES
not performed

8105 NE 140th Pl., Bothell WA 98011

b. SAFETY MARGIN



WASHINGTON STATE TOXICOLOGY LABORATORY
FORENSIC LABORATORY SERVICES BUREAU
WASHINGTON STATE PATROL
2203 AIRPORT WAY S, SUITE 360
SEATTLE WA 98134-2027
PHONE (206) 262-6100 FAX (206) 262-6145

ST

agency case #:
attn: Norman Thiersch
agency:
9509 29th Ave W
Everett WA 98204

date received: 4-2 -2008
date completed: 4-2 -2008

Last name | First name | Middle initial

sample container labeled: blood - peri: Y | urine: Y

BLOOD ETHANOL: > 0.03 g/100mL (99%)

BLOOD ANALYSES: not performed

B. STATUTORY/REGULATORY REQUIREMENTS – BLOOD TESTING: STANDARDS AS EVIDENCE

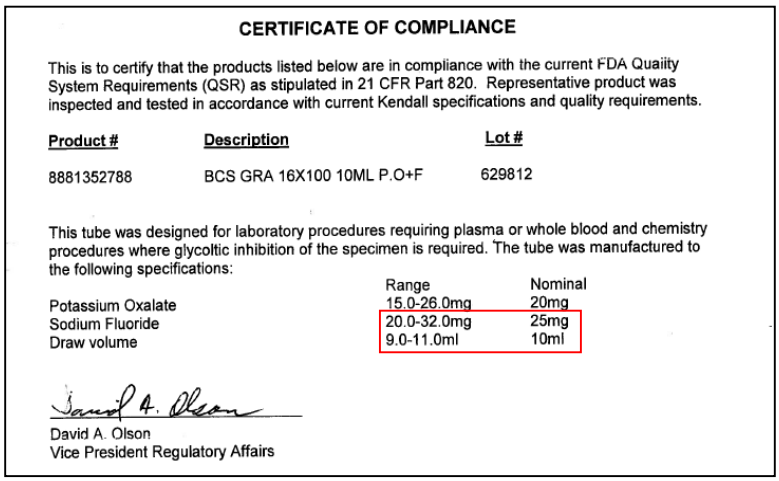
1. WAC 448-14-020: Operational discipline of blood samples for alcohol…(3) Sample container and preservative…Blood samples for alcohol analysis shall be preserved with…an enzyme poison *sufficient in amount to…stabilize the alcohol concentration.*

   a. A sufficient amount of preservative to stabilize alcohol concentration is not quantified.

   b. NCCLS, *Blood Alcohol Testing in the Clinical Laboratory; Approved Guideline*, T/DM6-A § 2.3.1 (1997).

      i. Defines sufficient amount to stabilize alcohol concentration as 10 mg/ml if the sample is not tested within 48 hours and not stored at -20°C.

      ii. 10 ml blood would require 100 mg preservative under standard.



CERTIFICATE OF COMPLIANCE

This is to certify that the products listed below are in compliance with the current FDA Quaiity System Requirements (QSR) as stipulated in 21 CFR Part 820. Representative product was inspected and tested in accordance with current Kendall specifications and quality requirements.

Product #	Description	Lot #
8881352788	BCS GRA 16X100 10ML P.O+F	629812

This tube was designed for laboratory procedures requiring plasma or whole blood and chemistry procedures where glycoltic inhibition of the specimen is required. The tube was manufactured to the following specifications:

	Range	Nominal
Potassium Oxalate	15.0-26.0mg	20mg
Sodium Fluoride	20.0-32.0mg	25mg
Draw volume	9.0-11.0ml	10ml

David A. Olson
Vice President Regulatory Affairs

      iii. If the blood sample fills this tube, is not tested within 48 hours and is not stored at -20°C, there is not sufficient amount of preservative present to stabilize alcohol concentration.

8105 NE 140th Pl., Bothell WA 98011

APPENDIX A

TOOLS & CONCEPTS
FOR
UNCERTAINTY AND RELIABILITY

8105 NE 140th Pl., Bothell WA 98011

Basic Measurement Concepts

Measurement: A set of empirical operations carried out to determine the quantity values that can reasonably be attributed to a quantity of interest. The objective of a measurement is to determine the value of the particular quantity being measured.

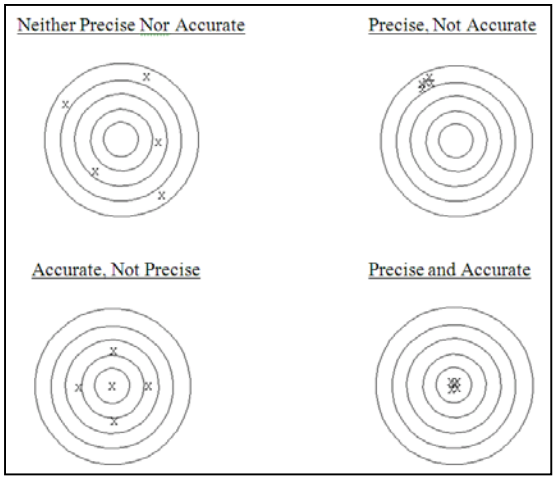
Measurand: The quantity intended to be measured.[448]

Direct Measurement: A measurement that senses the quantity of interest itself and maps it to a quantity value without the necessity of intermediate determinations.

Indirect Measurement: The determination of a quantity of interest through its relationship to other directly measured quantities.

Accuracy: The degree of agreement of a measured value with the "true" value of the quantity of interest. The degree of agreement expected from a measurement method/instrument is typically determined by comparing the mean of a set of measurements of a reference standard to the accepted value of the reference standard. Whether a measurement or instrument/method is deemed accurate is not an absolute judgment. Rather, accuracy is judged with respect to the use to be made of the data. What might be deemed accurate in one set of circumstances may not be accurate in another.

Precision: Precision is concerned with the variability or scatter of the individual results of replicate measurements. Measurements that are tightly grouped are considered precise while those with greater scatter are less so. As was the case with accuracy, precision is judged with respect to the use to be made of the data. What may be considered precise for one purpose may not be precise for another.



Measurement Interpretation – I: If a measurement value is to be interpretable, we must have an understanding of *how accurate* and *how precise* the measurement is. Absent such information, a measured value is simply a number, the meaning of which we know little about. Ideally, important measurements would be both accurate

[448] JCGM, *International Vocabulary of Metrology — Basic and General Concepts and Associated Terms (VIM)*, § 2.3 (2008).

and precise. That is, not only would such measurements yield mean values in close agreement with a "true" value, but individual values having a high degree of agreement with each other.

An objective characterization of accuracy and precision are necessary in order to determine the value of the particular quantity being measured. Such objective characterization can be supplied by statistics.

Basic Statistical Concepts

Population: The entire set or universe of objects sharing specific traits defining a class of objects.

Sample: A subset of objects selected from the population.

Distribution: The set of possible values of a random variable related through their frequency of occurrence or belief based relative likelihood.

Parameter: A characteristic of a population's distribution.

Statistic: A characteristic of a sample's distribution.

Descriptive Statistics: Utilizes data to describe the properties of a sample, not to make predictions based upon it.

Inferential Statistics: Utilizes data to draw inference or make predictions. A typical example is the use of sample data to generate a sample statistic from which an inference concerning a population parameter may be made.

Probability – Frequentist Interpretation: Probability is interpreted as relative frequency of occurrence over all sample data sets. As such, probabilities are objectively determined as a function of sampling data. Population parameters have unique, fixed true values that are unknown. The randomness lies in the sampling process, not the parameter. Since population parameters are nonrandom, probability statements cannot be made about their values. Nor can probability statements be made about a characteristic of a unique event. The parameter or characteristic either is or is not a particular value. The level of confidence associated with an inference refers to the confidence in the sampling/inferential process, *not* the actual quantity of interest. It tells us how often, over repeated samplings, our inference will happen to correspond to the true value.

Probability – Bayesian Interpretation: Probability is interpreted as an information-based "degree of belief" that an event will occur. Bayesian inference employs sampling data and any other information deemed relevant in the decision making process so that probability (degree of belief) may be based upon both objective and subjective components. In this framework, the parameters themselves are considered random so that probability statements can be made directly about their values. The same holds for a characteristic of a unique event. Thus, probability statements made concerning the value of a parameter or characteristic *are* about the actual quantity of interest. It tells us the probability that this particular inference is "true".

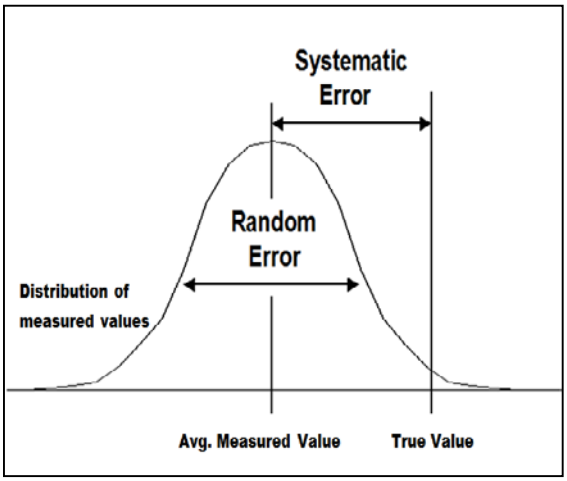
Measurement Error

Measurement Error: Traditionally, the quality of a measurement result was addressed through error analysis. This approach considered each measurand as having a unique true value. "The objective of measurement in the Error Approach is to determine an estimate of the true value that is as close as possible to that single true value. The deviation from the true value is composed of random and systematic errors."[449]

[449] JCGM, *International Vocabulary of Metrology — Basic and General Concepts and Associated Terms (VIM)*, § 0.1 (2008).

Systematic Error: The tendency of a set of measurements to consistently (on average) underestimate or overestimate the "true" value of the measurand by a given value or percentage. Most measurements have some amount of systematic error associated with them. Systematic error may be related to measuring methods, instruments or even empirically based calculations. It is a primary component of accuracy as it has a direct and regular impact on the degree of agreement of a measured value with the "true" value of the quantity of interest. Accordingly, "if a systematic error has not been accounted for, all [measured] values could be misleading."[450] Fortunately, once identified systematic error can be corrected for. "It is assumed that the result of a measurement has been corrected for all recognized significant systematic effects and that every effort has been made to identify such effects."[451]

Random Error: The unpredictable/random fluctuation in measurement results under fixed conditions. Random error is associated with precision. Unlike systematic error, random error cannot be corrected for. It is an inherent aspect of all measurement results. Although random error cannot be completely eliminated, it can be minimized by making a large number of measurements.



Arithmetic mean: This is a simple average of measurement values. It is determined by adding all measured values together and then dividing the sum by the number of values included in the sum. It is typically used when all measured values are considered to be equally reliable.

$$\bar{y} = \frac{1}{N} \cdot \sum_{i=1}^{N} y_i$$

Bias: Quantitative measure of systematic error. Bias is typically treated as either having a constant magnitude across a range of measured values or being proportional to the measured value obtained. When proportional, the bias is commonly reported as a percent bias. For chemical measurements, it is not uncommon for the bias to be proportional to measured values. "Whenever the true value of the measured quantity is needed…bias can be a serious problem."[452] Fortunately, once bias has been determined, systematic error can be easily

[450] Les Kirkup, *An Introduction to Uncertainty in Measurement* 33 (Cambridge University Press 2006).
[451] BIPM, *Evaluation of measurement data — Guide to the expression of uncertainty in measurement (GUM)*, § 3.2.4 (2008); NIST, *Guidelines for Evaluating and Expressing the Uncertainty of NIST Measurement Results*, NIST TN 1297 § 5.2 (1994).
[452] NIST, *NIST Special Publication 260-100*, 4 (1993).

accounted for. The bias of a method or instrument is ordinarily determined by comparing the mean of a set of measurements of a reference standard to its accepted value.

$$b_c = \bar{y} - Y_{ref}$$

$$b_{\%} = \frac{\bar{y} - Y_{ref}}{Y_{ref}}$$

Standard Deviation: Quantitative characterization of the variability/dispersion of individually measured values about their mean. The standard deviation is the root mean square deviation of measured values from their mean. Precision/random error is typically expressed in terms of a standard deviation. The determination of the standard deviation varies slightly depending on the source of our data. If the standard deviation has been determined from a population, we use what is commonly referred to as a population standard deviation. On the other hand, when our data comes from a sample, we use what is commonly referred to as a sample standard deviation. Throughout the remainder of this section the distinction will not be noted unless necessary but it is assumed that whenever employed, the correct standard deviation is utilized.

$$\sigma_{y_p} = \sqrt{\frac{1}{N} \cdot \sum_{i=1}^{N} (\bar{y} - y_i)^2}$$

$$\sigma_{y_s} = \sqrt{\frac{1}{n-1} \cdot \sum_{i=1}^{n} (\bar{y} - y_i)^2}$$

Coefficient of Variation: The standard deviation expressed as a proportion relative to the mean of a set of measurements. The coefficient of variation can be useful when combining standard deviations or comparing the variability of separate measurements.

$$\text{cv} = {}^{\sigma}/_{\bar{y}}$$

Bias Adjusted Mean/Best Estimate of True Value: The mean adjusted for bias. The bias adjusted mean is often considered the best estimate of the "true" value of the measurand. Whenever reporting the mean of a set of measurement, it should be corrected for bias. The correction applied depends upon whether the bias is constant or proportional.

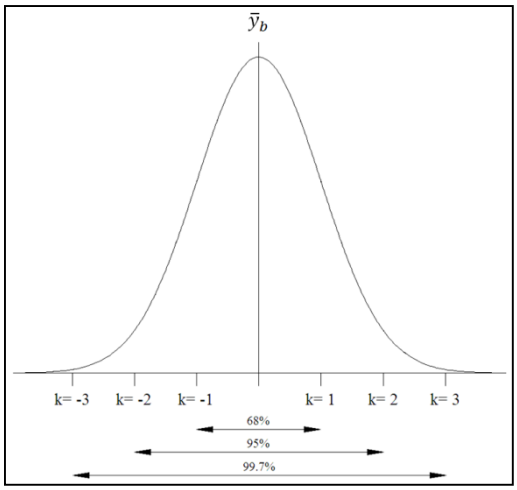
$$\bar{y}_b = \bar{y} - b_c$$

$$\bar{y}_b = \frac{\bar{y}}{1 + b_{\%}}$$

Confidence interval: A range of values symmetric about the bias adjusted mean constructed using a multiple of the standard deviation of the set of measurements and expected to cover the true value with a given level of confidence (likelihood).

$$C.I. = \bar{y}_b \pm k\sigma_y$$

The likelihood that the interval will overlap the true value is determined by the multiplier of the standard deviation (k), known as a coverage factor, and the underlying distribution. If the underlying distribution is Gaussian (normal) the likelihood associated with $k = 1, 2 \& 3$, is given in the following figure.



One should be very careful with the interpretation of a confidence interval. The focus of the level of confidence is not the true value. That is, the level of confidence does not refer to the probability that the true value lies within the interval. It either does or does not. Rather, the subject of the level of confidence is the sampling procedure. It tells you that based upon the procedure utilized, you will be able to construct an interval that will overlap the true value a given percent of the time. In technical terms, "[t]he confidence reflects the proportion of cases that the confidence interval would contain the true parameter value in a long series of repeated random samples under identical conditions."[453] The confidence interval is based upon frequentist philosophy and the existence of a singular true value.

Standard deviation (error) of the mean: Quantitative characterization of the variability/dispersion of sample means. Due to the Central Limit Theorem, the following relationship holds regardless of the underlying population distribution as long as the sample size is large enough.

$$\sigma_{\bar{y}} = \frac{\sigma_y}{\sqrt{N}}$$

Weighted mean: The weighted mean is an alternative way to determine the best estimate of the true value of a measurand. When combining multiple values determined for a given measurand, a weighted mean attaches more weight to those values considered more reliable.

$$\bar{y}_{wm} = \frac{\sum_{i=1}^{N} w_i \cdot y_i}{\sum_{i=1}^{N} w_i}$$

[453] ISO, *Statistics — Vocabulary and symbols — Part 1: General statistical terms and terms used in probability*, ISO 3534-1, § 1.28 (2007).

Traditional weighted mean: Frequently, the values to be combined are the arithmetic means from several sets of measurements. The traditional weighted mean relies upon the precision associated with each set of measurements to determine the weight to accord the mean associated with each set. The greater the precision associated with a given mean, the more confidence we have in the value, and the more weight it is accorded in combining the means to determine a best estimate of the true value. In this case the above expression becomes:

$$\bar{y}_{wm} = \frac{\sum_{i=1}^{N} \frac{n_i}{\sigma_i^2} . \bar{y}_i}{\sum_{i=1}^{N} \frac{n_i}{\sigma_i^2}}$$

The weighted mean should be employed when the values to be combined are not equally reliable.

Standard deviation of the Traditional Weighted Mean:

$$\sigma_{wm} = \frac{1}{\sqrt{\sum \frac{n_i}{\sigma_i^2}}}$$

Measurement Interpretation – II: If a measurement value is to be interpretable, we must have a *quantitative* determination of the *systematic* and *random error* associated with the measurement. Absent such information, a measured value is simply a number, the meaning of which we know little about. It has long been understood that no measurement result can be interpreted where only the value of the measurement itself is reported. Proper interpretation of a measured value requires knowledge and incorporation of the measurement's systematic and random error into any reported values.

Unfortunately, as useful as traditional error analysis is, "[i]t is now widely recognized that, when all of the known or suspected components of error have been evaluated and the appropriate corrections have been applied, there still remains an uncertainty about the correctness of the stated result, that is, a doubt about how well the result of the measurement represents the value of the quantity being measured."[454] Put simply, it is not possible to know the true value of a measurand or the error of a measurement result and hence how close a measurement result is to the true measurand value.[455]

Measurement Uncertainty

Measurement Uncertainty: "[F]or a given measurand and a given result of measurement of it, there is not one value but an infinite number of values dispersed about the result that are consistent with all of the observations and data and one's knowledge of the physical world, and that with varying degrees of credibility can be attributed to the measurand."[456] Accordingly, "[t]he objective of measurement in the Uncertainty Approach is not to determine a true value as closely as possible. Rather, it is assumed that the information from measurement only permits assignment of an interval of reasonable values to the measurand, based on the assumption that no mistakes have been made in performing the measurement."[457]

---

[454] BIPM, *Evaluation of measurement data — Guide to the expression of uncertainty in measurement (GUM)*, § 0.2, 3.2.2 – 3.2.3 (2008).
[455] BIPM, *Evaluation of measurement data — Guide to the expression of uncertainty in measurement (GUM)*, §3.2.1, 3.3.2 (2008); Ehrlich, *Evolution of philosophy and description of measurement* 12 ACCRED. QUAL. ASSUR. 201, 210 (2007).
[456] BIPM, *Evaluation of measurement data — Guide to the expression of uncertainty in measurement (GUM)*, § 5.2 (2008).
[457] JCGM, *International Vocabulary of Metrology — Basic and General Concepts and Associated Terms (VIM)*, § 0.1 (2008).

Contrary to the traditional approach, then, the measurand is not treated as having a unique "true" value. Instead, the measurand is deemed to consist of a set of "true" values. Measurement uncertainty is a quantitative statement characterizing the dispersion of values that can actually and reasonably be attributed "to a measurand based on the information available including systematic and random effects…and any other factors that may impact the measurement or test process or result."[458] Measurement uncertainty is based upon the Bayesian notion of probability as a measure of degree of belief.

Standard Uncertainty: The total uncertainty associated with any measurement result is typically the result of the combination of several smaller uncertainties associated with particular aspects of the measurement process. Each component of uncertainty that contributes to the uncertainty of a measurement result is known as a standard uncertainty. Each standard uncertainty is expressed and treated as, and may in fact be, a standard deviation.

$$\mu \equiv \sigma$$

Relative Standard Uncertainty: The standard uncertainty expressed as a proportion relative to the mean of a set of measurements. It can be useful when combining standard uncertainties or comparing the uncertainty of separate measurements.

$$\mu_r = \frac{\mu_y}{|\bar{y}_b|}$$

Type A Uncertainty: Component of uncertainty that has been determined by the statistical analysis of measured values. Determination is based on frequency distributions and any statistically valid method for data analysis. An example is the standard deviation determined from a set of measurements.

Type B Uncertainty: Component of uncertainty that has been determined by means other than the statistical analysis of measured values. Determination assumes *a priori* distributions based on relevant information and scientific judgment. Examples include information provided by instrument manufacturer, metrological certifications and reference publications.

Combined Uncertainty: The combination of all the standard uncertainties associated with a measurement. The individual standard uncertainties are combined in the same manner as standard deviations. Assuming the standard uncertainties are random and independent, the combined uncertainty is the root sum square of the standard uncertainties. The combined uncertainty is expressed and treated as, and may in fact be, a standard deviation.

$$\mu_c = \sqrt{\sum_{i=1}^{n} \mu_i^2}$$

When determining the combined uncertainty of a measurement it is critical to include all significant components of uncertainty. Failure to do so will cause an underestimate of the uncertainty misleading others to believe that the result is more precise than it actually is.

[458] JCGM, *International Vocabulary of Metrology — Basic and General Concepts and Associated Terms (VIM)*, § 2.26 (2008); ASTM, *Standard Terminology Relating to Quality and Statistics*, E 456 § 3 (2008).

Expanded Uncertainty: Obtained by multiplying the combined uncertainty by a coverage factor.

$$U = k\mu_c$$

A coverage factor is chosen such that when the expanded uncertainty is expressed as part of a complete measurement result it conveys a range of values that can actually and reasonably be attributed to a measurand with a given level of confidence. The level of confidence associated with a given coverage factor is determined by the measurement's underlying distribution. If the underlying distribution is Gaussian (normal) the level of confidence associated with $k = 1.64, 1.96$ & 2.576, is given in the following table.

k	level of confidence
1.64	90%
1.96	95%
2.576	99%

Measurement Result: "In general, the result of a measurement is only an approximation or estimate of the value of the specific quantity subject to measurement, that is, the measurand, and thus the result is complete only when accompanied by a quantitative statement of its uncertainty."[459] Moreover, "[i]t is assumed that the result of a measurement has been corrected for all recognized significant systematic effects and that every effort has been made to identify such effects."[460] Accordingly, a complete measurement result consists of the best estimate of the true value of the measurand, typically the bias adjusted mean, accompanied by the expanded uncertainty and its associated level of confidence.

$$Y = \bar{y}_b \pm U\ (99\%)$$

This is interpreted to mean that the best estimate of the value attributable to the measurand Y is $\bar{y}_b$, and that $\bar{y}_b - U$ to $\bar{y}_b + U$ is the range of values that could actually be attributed to Y with a 99% level of confidence.

Coverage Interval: An "interval containing the set of true quantity values of a measurand with a stated probability, based on the information available."[461] Ordinarily the coverage interval is derived from the expanded uncertainty and is symmetric about the mean so that it can be expressed as:

$$\mathrm{C} = \bar{y}_b \pm U\ (99\%)$$

Note that the coverage interval is identical to the measurement result. Unlike the confidence interval, the coverage interval is based upon Bayesian philosophy so that it refers directly to the quantity of interest, the "true" value of the measurand. In this context, the level of confidence is the probability, understood as a degree of belief, "that the set of true quantity values of a measurand is contained within a specified coverage interval."[462] It should also be noted that the coverage interval need not be symmetric about the mean.

[459] NIST *Guidelines for Evaluating and Expressing the Uncertainty of NIST Measurement Results*, NIST TN 1297 § 2.1 (1994); BIPM, *Evaluation of measurement data — Guide to the expression of uncertainty in measurement (GUM)*, § 7.1.4 (2008).
[460] BIPM, *Evaluation of measurement data — Guide to the expression of uncertainty in measurement (GUM)*, § 3.2.4 (2008); NIST, *Guidelines for Evaluating and Expressing the Uncertainty of NIST Measurement Results*, NIST TN 1297 § 5.2 (1994).
[461] JCGM, *International Vocabulary of Metrology — Basic and General Concepts and Associated Terms (VIM)*, § 2.36 (2008).
[462] JCGM, *International Vocabulary of Metrology — Basic and General Concepts and Associated Terms (VIM)*, § 2.37 (2008).

Measurement Interpretation – III: For even the most carefully performed measurement, a unique "true" value for a measurand can never be determined. All that can ever be given is a set of values, all of which may actually and reasonably be assigned as "true" values. If a measurement value is to be interpretable, it must be *corrected for bias* and accompanied by a *quantitative estimate of its uncertainty*. Absent such information, a measured value is simply a number, the meaning of which we know little about.

> "Knowledge of the uncertainty associated with measurement results is essential to the interpretation of the results. Without quantitative assessments of uncertainty, it is impossible to decide whether observed differences between results reflect more than experimental variability, whether test items comply with specifications, or whether laws based on limits have been broken. Without information on uncertainty, there is a risk of misinterpretation of results. Incorrect decisions taken on such a basis may result in unnecessary expenditure in industry, incorrect prosecution in law, or adverse health or social consequences."[463]

Reliability: While in quantitative measurement, the quality of the measurement is determined by its associated uncertainty, the quality of nonquantitative tests is determined by measures of their reliability.

Traditional Measures of Reliability: Some traditional measures of reliability (see below) can be defined with respect to the following two-way table:

	Test Result A	Test Result ¬A	
Condition A	True Positive N_{TP}	False Negative (Type I error) N_{FN}	$N_{TP} + N_{FN}$
Condition ¬A	False Positive (Type II error) N_{FP}	True Negative N_{TN}	$N_{FP} + N_{TN}$
	$N_{TP} + N_{FP}$	$N_{FN} + N_{TN}$	N

False negative (Type I error) rate: Percent rejection of true condition.

$$FNR = [N_{FN} / (N_{TP} + N_{FN})]$$

False positive (Type II error) rate: Percent failure to reject false condition.

$$FPR = [N_{FP} / (N_{FP} + N_{TN})]$$

Sensitivity: Percent confirming a true condition.

$$S_e = [N_{TP} / (N_{TP} + N_{FN})]$$

Specificity: Percent rejecting a false condition.

[463] ISO, *Guidance for the use of repeatability, reproducibility and trueness estimates in measurement uncertainty estimation*, ISO/TS 21748 DRAFT REVISION, v (2009); ISO, *Guidance for the use of repeatability, reproducibility and trueness estimates in measurement uncertainty estimation*, ISO/TS 21748, v (2004).

$$S_p = [N_{TN} /(N_{FP} + N_{TN})]$$

Positive predictive value: Percent indicating condition true that are correct.

$$P_{pv} = [N_{TP} /(N_{FP} + N_{TP})]$$

Negative predictive value: Percent indicating condition false that are correct.

$$N_{pv} = [N_{TN} /(N_{FN} + N_{TN})]$$

BAYES THEOREM: States that the probability of a hypothesis being true given some result is proportional to the probability of the hypothesis being true prior to obtaining the result multiplied by the probability of obtaining the result assuming the hypothesis is true. Bayes Theorem provides a rigorous means of incorporating prior information into a measurement. It can be written as:

$$p(H \mid I) \propto p(I|H)p(H)$$

where

$p(H \mid I)$ = Posterior probability: Probability of H given result I.
$p(H)$ = Prior probability: Independent probability of H prior to result I.
$p(I \mid H)$ = Probability of result I if H true.

Likelihood Ratio: Using Bayes theorem, the meaning of a test result can be judged by the likelihood ratio obtained. The likelihood ratio is defined as:

$$L(I \mid H) = \frac{p(I \mid H)}{p(I \mid \neg H)}.$$

This is a measure of the impact of the test result on the likelihood of H, that is of how much the test result has increased or decreased the pretest likelihood of H. Th

Functional Relationships, Measurement Functions and Propagation of Uncertainty

Algorithmic Determinations: When the quantity of interest cannot be measured directly, we must rely upon mathematical relationships between the quantity of interest and other measured and/or "given" values to calculate the quantity of interest. Each measured value and many "given" values have uncertainty associated with them. These uncertainties propagate through the calculation and are imparted to the value determined for the quantity of interest.

Measurement Function – General Form: A functional relationship between the quantity of interest and the input quantities (measured and/or "given" values) needed to calculate it.

$$Y = f(X, W \cdots Z)$$

Best Estimate of True Value – General Form: The best estimate of a quantity value based on a measurement function is given by plugging in the best estimate for each of the input quantities

$$Y_b = f(x_b, w_b \cdots z_b)$$

Propagation of Uncertainty – General Form: For a quantity value based upon a general measurement function:

All Circumstances

$$\mu_y = \sqrt{\sum_{i=1}^{N}\left(\frac{\partial f}{\partial x_i}\cdot\mu_{x_i}\right)^2 + 2\sum_{i=1}^{N-1}\sum_{j=i+1}^{N}\frac{\partial f}{\partial x_i}\cdot\frac{\partial f}{\partial x_j}\cdot\mu_{x_i x_j}}$$

$$\mu_y = \sqrt{\left(\left.\frac{\partial f(X,Z)}{\partial X}\right|_{x_b z_b}\cdot\mu_x\right)^2 + \left(\left.\frac{\partial f(X,Z)}{\partial Z}\right|_{x_b z_b}\cdot\mu_z\right)^2 + 2\cdot\left.\frac{\partial f(X,Z)}{\partial X}\right|_{x_b z_b}\cdot\left.\frac{\partial f(X,Z)}{\partial Z}\right|_{x_b z_b}\cdot\mu_{xz}}$$

$$\mu_y \le \left|\left.\frac{\partial f(X,\cdots,Z)}{\partial X}\right|_{x_b\cdots z_b}\cdot\mu_x + \cdots + \left|\left.\frac{\partial f(X,\cdots,Z)}{\partial Z}\right|_{x_b\cdots z_b}\cdot\mu_z\right.\right.$$

Independent Input Quantities

$$\mu_y = \sqrt{\sum_{i=1}^{N}\left(\frac{\partial f}{\partial x_i}\cdot\mu_{x_i}\right)^2}$$

$$\mu_y = \sqrt{\left(\left.\frac{\partial f(X,\cdots,Z)}{\partial X}\right|_{x_b\cdots z_b}\cdot\mu_x\right)^2 + \cdots + \left(\left.\frac{\partial f(X,\cdots,Z)}{\partial Z}\right|_{x_b\cdots z_b}\cdot\mu_z\right)^2}$$

Covariance: A measure of the association between two random variables. If two input quantities are independent then the covariance will be zero. When two input quantities are not independent this term appears in the propagation of uncertainty calculation to account for the dependence.

$$\mu_{xy} = \frac{1}{N}\cdot\sum_{i=1}^{N}(\bar{x}-x_i)(\bar{y}-y_i)$$

Measurement Function – Measured quantity multiplied by a constant:

$$Y = a\cdot X$$

$$Y_b = a\cdot x_b$$

$$\mu_y = a\cdot\mu_x$$

Measurement Function – Variable raised to a constant power:

$$Y = X^n$$

$$Y_b = x_b^n$$

$$\mu_{r_y} = \frac{\mu_y}{|Y_b|} = |n| \frac{\mu_x}{|x_b|}$$

Measurement Function – Sums and differences:

$$Y = X - W + \cdots + Z$$

$$Y_b = x_b - w_b + \cdots + z_b$$

Independent

$$\mu_y = \sqrt{\mu_x^2 + \mu_w^2 + \cdots + \mu_z^2}$$

All Circumstances

$$\mu_y \leq \mu_x + \mu_w + \cdots + \mu_z$$

Measurement Function – Products and quotients:

$$Y = \frac{X \times \cdots \times W}{Z \times \cdots \times Q}$$

$$Y_b = \frac{x_b \times \cdots \times w_b}{z_b \times \cdots \times q_b}$$

Independent

$$\mu_{r_y} = \frac{\mu_y}{|Y_b|} = \sqrt{\left(\frac{\mu_x}{x_b}\right)^2 + \left(\frac{\mu_w}{w_b}\right)^2 + \cdots + \left(\frac{\mu_z}{z_b}\right)^2 + \left(\frac{\mu_q}{q_b}\right)^2}$$

All Circumstances

$$\mu_{r_y} = \frac{\mu_y}{|Y_b|} \leq \frac{\mu_x}{|x_b|} + \frac{\mu_w}{|w_b|} + \cdots + \frac{\mu_z}{|z_b|} + \frac{\mu_q}{|q_b|}$$

Examples: BAC Results & Calculations

Breath Testing: Like any other measurement, forensic breath alcohol concentration tests have both bias and uncertainty associated with them. Both need to be determined and incorporated into a complete test result.

Best estimate for "true" BrAC (Bias adjusted mean): $(\overline{BrAC_b})$
Best estimate for "true" BrAC determined by computing the bias adjusted mean.

Machine bias: (b_M)
Determined during calibration and which will be deemed proportional to the concentration measured.

Interferent bias: (b_I)

Most breath test instruments are designed to detect the presence of interferents on an individual's breath. However, some are programmed such that they will only do so if the interferent exceeds a particular level. There are several ways one might try to determine the average impact/bias such intereferent will have on a breath test below the level of detection but which will nonetheless contribute to the reported value. One could consult the literature to determine if there are published values. Another way is to postulate an underlying distribution based upon all the known information and determine the mean (expected) contribution due to bias based on the distribution. The bias due to this source will be a constant offset.

Best estimate for "true" BrAC (Bias adjusted mean):

$$\overline{BrAC}_b = \frac{\overline{BrAC}}{1 + b_M} - b_I$$

Combined uncertainty for BAC based on measurement:[464] (μ_{BrAC})
There are several sources of uncertainty that may be associated with a breath test. The ones considered here comprise only a subset and may or may not be relevant to your test. For ease of illustration they are treated as being independent.

Reference material: (μ_r)
The standard uncertainty associated with the reference material utilized to calibrate machine.

Machine precision: (μ_σ)
The precision of the breath test machine determined at the time of it's calibration and expressed as a standard uncertainty.

Bias: (μ_b)
The standard uncertainty associated with the value determined for the bias.

Sampling: (μ_s)
The standard uncertainty due to circumstances arising during the collection of breath samples.

Combined uncertainty for BAC based on measurement:

$$\mu_{c_{BrAC}} = \overline{BrAC}_b \cdot \sqrt{\left(\frac{\mu_r}{r_b}\right)^2 + \left(\frac{\mu_\sigma}{\sigma_b}\right)^2 + \left(\frac{\mu_b}{b_b}\right)^2 + \left(\frac{\mu_s}{S_b}\right)^2}$$

Complete Result:

$$BrAC = \overline{BrAC}_b \pm k\mu_{c_{BrAC}}$$

[464] The example assumes that the uncertainties are independent.

Breath as an indirect measure of blood: When the breath alcohol concentration is being utilized as an indirect measure of blood alcohol concentration, the breath result must be converted into one for blood. This involves a conversion factor (M) between breath and blood alcohol concentration which the literature illustrates has a great deal of uncertainty associated with it. This uncertainty must also be factored into a reported result.

Functional Relationship:

$$BAC = M \cdot BrAC$$

Combined Uncertainty:

$$\mu_{c_{BAC}} = \sqrt{\left(\frac{\partial BAC(M,BrAC)}{\partial M} \cdot \mu_M\right)^2 + \left(\frac{\partial BAC(M,BrAC)}{\partial BrAC} \cdot \mu_{BrAC}\right)^2 + 2 \cdot \frac{\partial BAC(M,BrAC)}{\partial M} \cdot \frac{\partial BAC(M,BrAC)}{\partial BrAC} \cdot \mu_{M,BrAC}}$$

$$= \sqrt{(\overline{BrAC}_b \cdot \mu_M)^2 + (M_b \cdot \mu_{BrAC})^2 + 2 \cdot \overline{BrAC}_b \cdot M_b \cdot \mu_{M,BrAC}}$$

Complete Result:

$$BAC_b = M_b \cdot \overline{BrAC}_b \pm k\mu_{c_{BAC}}$$

Widmark's Formula: For the determination of blood alcohol content based on the number of drinks consumed.

Functional Relationship 1:[465] Assuming post-absorptive.

$$C_t = \frac{NdZ}{W\tau} - \beta t$$

The Variables:

$C_t \equiv$ BAC at time t
$N \equiv$ Number of drinks
$d \equiv$ Density of alcohol
$Z \equiv$ Ethanol per drink
$W \equiv$ Body weight
$\tau \equiv$ Volume of distribution
$\beta \equiv$ Alcohol elimination rate
$t \equiv$ Time since drinking began

Combined Uncertainty:[466]

$$\mu_{C_t} = \sqrt{\left[\sum \left(\frac{\partial C_t}{\partial x_i} \cdot \mu_{x_i}\right)^2\right] + 2 \cdot \frac{\partial C_t}{\partial \tau} \cdot \frac{\partial C_t}{\partial \beta} \cdot \mu_{\tau,\beta}}$$

[465] This example is taken from Gullberg, *Estimating the uncertainty associated with Widmark's equation as commonly applied in forensic toxicology* 172 FOR. SCI. INT. 33 (2007), and fills in some blanks not explicitly shown therein.
[466] Assuming only τ and β are not independent.

$$= \sqrt{\left(\frac{dZ}{W\tau}\cdot\mu_N\right)^2 + \left(\frac{NZ}{W\tau}\cdot\mu_d\right)^2 + \left(\frac{Nd}{W\tau}\cdot\mu_Z\right)^2 + \left(-\frac{NdZ}{W^2\tau}\cdot\mu_W\right)^2 + \left(-\frac{NdZ}{W\tau^2}\cdot\mu_\tau\right)^2 + (-t\cdot\mu_\beta)^2 + (-\beta\cdot\mu_t)^2 + 2\cdot\left(-\frac{NdZ}{W\tau^2}\right)(-t)\cdot\mu_{\tau,\beta}}$$

Complete Result:

$$C_{t_b} = \frac{N_b d_b Z_b}{W_b \tau_b} - \beta_b t_b \pm k\mu_{C_t}$$

Functional Relationship 2: Accounting for rate of absorption.

$$C_t = \frac{NdZ(1-e^{-\gamma t})}{W\tau} - \beta t$$

The Variables:

- $C_t \equiv$ BAC at time t
- $N \equiv$ Number of drinks
- $d \equiv$ Density of alcohol
- $Z \equiv$ Ethanol per drink
- $\gamma \equiv$ Alcohol absorption rate
- $W \equiv$ Body weight
- $\tau \equiv$ Volume of distribution
- $\beta \equiv$ Alcohol elimination rate
- $t \equiv$ Time since drinking began

Combined Uncertainty:[467]

$$\mu_{C_t} = \sqrt{\left[\sum\left(\frac{\partial C_t}{\partial x_i}\cdot\mu_{x_i}\right)^2\right] + 2\cdot\frac{\partial C_t}{\partial\tau}\cdot\frac{\partial C_t}{\partial\beta}\cdot\mu_{\tau,\beta}}$$

$$= \sqrt{\left(\frac{dZ(1-e^{-\gamma t})}{W\tau}\cdot\mu_N\right)^2 + \left(\frac{NZ(1-e^{-\gamma t})}{W\tau}\cdot\mu_d\right)^2 + \left(\frac{Nd(1-e^{-\gamma t})}{W\tau}\cdot\mu_Z\right)^2 + \left(-\frac{NdZ(1-e^{-\gamma t})}{W^2\tau}\cdot\mu_W\right)^2 + \left(-\frac{NdZ(1-e^{-\gamma t})}{W\tau^2}\cdot\mu_\tau\right)^2 + \left(\frac{NdZte^{-\gamma t}}{W\tau}\cdot\mu_\gamma\right)^2 + (-t\cdot\mu_\beta)^2 + \left(\left(\frac{NdZ\gamma e^{-\gamma t}}{W\tau} - \beta\right)\cdot\mu_t\right)^2 + 2\cdot\left(-\frac{NdZ(1-e^{-\gamma t})}{W\tau^2}\right)(-t)\cdot\mu_{\tau,\beta}}$$

Complete Result:

$$C_{t_b} = \frac{N_b d_b Z_b (1-e^{-\gamma_b t_b})}{W_b \tau_b} - \beta_b t_b \pm k\mu_{C_t}$$

[467] Assuming only τ and β are not independent.

APPENDIX B

MISCELLANEOUS RESOURCES

I. ACRONYMS

A2LA	American Association of Laboratory Accreditation
AAFS	American Academy of Forensic Sciences
ANSI	American National Standards Institute
ASCLD	American Society of Crime Lab Directors
ASQ	American Society for Quality
ASTM	American Society for Testing and Materials
BIPM	International Bureau of Weights and Measures
FQS-I	Forensic Quality Services - International
IEC	International Electrotechnical Commission
IEEE	Institute of Electrical and Electronics Engineers
ILAC	International Laboratory Accreditation Cooperation
ISO	Greek for equal. Not an acronym.
IUPAC	International Union of Pure and Applied Chemistry
JCGM	Joint Committee for Guides in Metrology
NAS	National Academy of Sciences
NIST	National Institute of Standards and Technology
NIST-OLES	National Institute of Standards and Technology Law Enforcement Standards Office
NFSTC	National Forensic Science Technology Center
OIML	International Organization of Legal Metrology
SOFT	Society of Forensic Toxicologists
SWGDRUG	Scientific Working Group for the Analysis of Seized Drugs
SWGTOX	Scientific Working Group for Forensic Toxicology
UNODC	United Nations Office on Drugs and Crime
WADA	World Anti–Doping Agency
WMO	World Metrological Organization

II. METROLOGY INSTITUTES

- International Bureau of Weights and Measures (BIPM): http://www.bipm.org/
- International Organization of Legal Metrology (OIML) http://www.oiml.org/en
- National Institute of Standards and Technology (NIST) http://www.nist.gov/
- World Metrological Organization (WMO) http://www.wmo.int/pages/index_en.html
- List – National Metrology Institutes (NMIs) http://www.bipm.org/en/cipm-mra/participation/signatories.html
- List – State Metrology Labs/Offices – List http://www.nist.gov/pml/wmd/labmetrology/lab-contacts-ac.cfm

III. STANDARDS ORGANIZATIONS

- American National Standards Institute (ANSI)
  http://www.ansi.org/
- American Society for Testing and Materials (ASTM)
  http://www.astm.org/
- Eurachem
  http://www.eurachem.org/
- International Organization for Standardization (ISO)
  http://www.iso.org/iso/home.htm
- International Organization of Legal Metrology (OIML)
  http://www.oiml.org/en
- International Union of Pure and Applied Chemistry (IUPAC)
  http://www.iupac.org/
- International Laboratory Accreditation Cooperation (ILAC)
  https://www.ilac.org/
- International Electrotechnical Commission (IEC)
  http://www.iec.ch/
- National Institute of Standards and Technology Law Enforcement Standards Office (NIST-OLES)
  http://www.nist.gov/oles/
- Scientific Working Group for the Analysis of Seized Drugs (SWGDRUG)
  http://www.swgdrug.org/
- Scientific Working Group for Forensic Toxicology (SWGTOX)
  http://www.swgtox.org/
- United Nations Office on Drugs and Crime (UNODC)
  http://www.unodc.org/unodc/index.html
- World Anti–Doping Agency (WADA)
  http://www.wada-ama.org/

IV. SELECT STANDARDS

Terminology

- Joint Committee for Guides in Metrology, *International Vocabulary of Metrology — Basic and General Concepts and Associated Terms (VIM)* JCGM 200 (2008)
  http://www.bipm.org/utils/common/documents/jcgm/JCGM_200_2008.pdf

- Eurachem, *Terminology in Analytical Measurement – Introduction to VIM 3* (*TAM*) (2011)
  http://www.accredia.it/UploadDocs/1629_TAM_2011_Final_web.pdf

- American Society for Testing and Materials, *Standard Terminology Relating to Forensic Science*, E 1732 (2005).

http://www.astm.org/Standards/E1732.htm (purchase required).

Quantities and Units

- International Organization for Standardization, *Quantities and Units*, ISO 80000 Parts 1 – 14 [*Part* 1: *General*; *Part* 2: *Mathematical signs and symbols to be used in the natural sciences and technology*; *Part* 3: *Space and time*; *Part* 4: *Mechanics*; *Part* 5: *Thermodynamics*; *Part* 6: *Electromagnetism; Part* 7: *Light*; *Part* 8: *Acoustics*; *Part* 9: *Physical chemistry and molecular physics*; *Part* 10: *Atomic and nuclear physics*; *Part* 11: *Characteristic numbers*; *Part* 12: *Solid state physics*; *Part* 13: *Information science and technology*; *Part* 14: *Telebiometrics related to human physiology*.](2009)
  Available from ISO (purchase required)

- International Bureau of Weights and Measures, *The International System of Units (SI)* (8th ed. 2008)
  http://www.bipm.org/utils/common/pdf/si_brochure_8.pdf

- National Institute of Standards and Technology, *Guide for the Use of the International System of Units (SI)*, NIST SP 811 (2008)
  http://physics.nist.gov/cuu/pdf/sp811.pdf

Traceability

- International Union of Pure and Applied Chemistry, *Metrological traceability of measurement results in chemistry: Concepts and implementation*, IUPAC Technical Report (2011)
  http://pac.iupac.org/publications/pac/pdf/2011/pdf/8310x1873.pdf.

- Eurachem, *Traceability in Chemical Measurement: A Guide to Achieving Comparable Results* (2003)
  http://www.eurachem.org/images/stories/Guides/pdf/EC_Trace_2003.pdf

Validation

- International Union of Pure and Applied Chemistry, *Harmonized Guidelines for Single Laboratory Validation of Methods of Analysis*, IUPAC Technical Report 74(5) PURE APPL. CHEM. 835 (2002)
  http://www.iupac.org/publications/pac/2002/pdf/7405x0835.pdf

- Eurachem, *The Fitness for Purpose of Analytical Methods: A Laboratory Guide to Method Validation and Related Topics* (1998)
  http://www.gnbsgy.org/PDF/Eurachem%20Guide%20Validation[1].pdf

- United Nations Office on Drugs and Crime, *Guidance for the Validation of Analytical Methodology and Calibration of Equipment used for Testing of Illicit Drugs in Seized Materials and Biological Specimens* ST/NAR/41 (1995)
  http://www.unodc.org/documents/scientific/validation_E.pdf

- Scientific Working Group for Forensic Toxicology, *Standard Practices for Method Validation in Forensic Toxicology* (2013)

http://www.swgtox.org/documents/Validation3.pdf

Good Measurement Practices

- International Organization for Standardization, *General requirements for the competence of testing and calibration laboratories*, ISO 17025 (2005)
  http://www.iso.org/iso/catalogue_detail?csnumber=39883 (purchase required)

- National Institute of Standards and Technology, *National Voluntary Laboratory Accreditation Program Procedures and General Requirements*, NIST HB150 (2006)
  http://www.nist.gov/nvlap/upload/nist-handbook-150.pdf (contains the technical requirements of ISO 17025 in sections 4 and 5)

- National Institute of Standards and Technology, *Selected Laboratory and Measurement Practices, and Procedures, to Support Basic Mass Calibrations*, NIST IR 6969 (2012)
  http://www.nist.gov/pml/wmd/pubs/upload/NISTIR_6969_Feb2012.pdf

- World Anti–Doping Agency, *International Standard for Laboratories* (2012)
  http://www.wada-ama.org/Documents/World_Anti-Doping_Program/WADP-IS-Laboratories/ISL/WADA_Int_Standard_Laboratories_2012_EN.pdf

- International Laboratory Accreditation Cooperation, *Guideline for Forensic Science Laboratories*, ILAC G19 (2002)
  https://www.ilac.org/documents/g19_2002.pdf

- Scientific Working Group for the Analysis of Seized Drugs, ***SWGDRUG*** *Recommendations of the Scientific Working Group for the Analysis of Seized Drugs* (2013)
  http://www.swgdrug.org/Documents/SWGDRUG%20Recommendations%20Version%206-1.pdf

- United Nations Office on Drugs and Crime, *Recommended Guidelines for Quality Assurance and Good Laboratory Practice* ST/NAR/25 (1995)
  http://www.unodc.org/unodc/en/scientists/publications_manuals.html

- Society of Forensic Toxicologists, *Forensic Toxicology Laboratory Guidelines* (2006)
  http://www.soft-tox.org/files/Guidelines_2006_Final.pdf

Calibration

- International Organization for Standardization, *Linear calibration using reference materials*, ISO 11095 (1996)
  http://www.iso.org/iso/catalogue_detail.htm?csnumber=1060 (purchase required)

- International Laboratory Accreditation Cooperation, *Guidelines for the determination of calibration intervals of measuring instruments*, ILAC G24 (2007)
  https://www.ilac.org/documents/ILAC_G24_2007.pdf

Reference Materials

- International Organization for Standardization, *General Requirements for the Competence of Reference Material Producers*, ISO Guide 34:2009 (2009)
  http://www.iso.org/iso/catalogue_detail.htm?csnumber=50174 (purchase required)

- International Organization for Standardization, *Reference Materials—General and Statistical Principles for Certification*, ISO Guide 35:2006(E) (2006)
  http://www.iso.org/iso/catalogue_detail.htm?csnumber=39269 (purchase required)

- National Institute of Standards and Technology, *Standard Reference Materials: Statistical Aspects of the Certification of Chemical Batch SRMs*, NIST SP 260-125 (1996)
  http://www.nist.gov/srm/upload/SP260-125.PDF

- National Institute of Standards and Technology, *Standard Reference Materials: Handbook for SRM Users*, NIST SP 260-100 (1993)
  http://www.nist.gov/srm/upload/SP260-100.pdf

- International Laboratory Accreditation Cooperation, *Guidelines for the Selection and Use of Reference Materials*, ILAC G9 (2005)
  https://www.ilac.org/documents/ILAC_G9_2005_guidelines_for_the_selection_and_use_of_reference_material.pdf

Uncertainty/Reporting results

- Joint Committee for Guides in Metrology, *Evaluation of measurement data — Guide to the expression of uncertainty in measurement* (*GUM*), JCGM 100 (2008)
  http://www.bipm.org/utils/common/documents/jcgm/JCGM_100_2008_E.pdf

- National Institute of Standards and Technology, *Guidelines for Evaluating and Expressing the Uncertainty of NIST Measurement Results*, NIST TN 1297 (1994)
  http://physics.nist.gov/Pubs/guidelines/TN1297/tn1297s.pdf

- Eurachem, *Quantifying Uncertainty in Analytical Measurement*, CG-4 (2000)
  http://www.eurachem.org/guides/pdf/QUAM2000-1.pdf

- Eurachem, *Measurement uncertainty arising from sampling: A guide to methods and approaches*, (2007)
  http://www.eurachem.org/guides/pdf/UfS_2007.pdf

- International Organization for Standardization, *Guidance for the use of repeatability, reproducibility and trueness estimates in measurement uncertainty estimation*, ISO 21748 (2010)
  http://www.iso.org/iso/home/store/catalogue_tc/catalogue_detail.htm?csnumber=46373 (purchase required)

- American Society of Crime Laboratory Directors/Laboratory Accreditation Board, *Guidance on the Estimation of Measurement Uncertainty – Annex A: Details on the NIST 8-Step Process* (2011) http://www.ascld-lab.org/documents/AL-PD-3062.pdf

- American Society for Testing and Materials, *Standard Practice for Reporting Opinions of Scientific or Technical Experts*, E-620 (2011). http://www.astm.org/search/fullsite-search.html?query=e-620& (purchase required)

V. ACCREDITATION BODIES

- ANSI-ASQ National Accreditation Board http://www.anab.org/

- American Association of Laboratory Accreditation (A2LA) http://www.a2la.org/

- American Society of Crime Laboratory Directors/Laboratory Accreditation Board (ASCLD/LAB) http://www.ascld-lab.org/

- International Laboratory Accreditation Cooperation (ILAC) https://www.ilac.org/

- NIST, National Voluntary Laboratory Accreditation Program (NVLAP) http://www.nist.gov/nvlap/

VI. Books

- Brach, R., & Dunn, P., *Uncertainty Analysis for Forensic Science.* (2nd ed. 2010).

- Bucher, J. (Ed.). *The Metrology Handbook*, (2004).

- Dieck, R., *Measurement Uncertainty Methods and Applications*, (4th Ed. 2007).

- Gullberg, R., *Statistical Applications in Forensic Toxicology*, Medical-Legal Aspects of Alcohol, Ch. 18 (James Garriott ed., 5th ed. 2009).

- Howson, *Scientific Reasoning The Bayesian Approach* (Open Court 2006).

- Kirkup, L. *An Introduction to Uncertainty in Measurement* (2006).

- National Research Council of the National Academy of Sciences, *Strengthening Forensic Science in the United States: A Path Forward* (2009).

- Taylor, *An Introduction to Error Analysis: The Study of Uncertainties in Physical Measurements* (2nd Ed. 1997).

- Vosk, T., and Emery, A., *Forensic Metrology: Scientific Measurement and Inference for Lawyers, Judges, and Criminalists* (CRC/Taylor Francis Group – Available Sept. 2014)

- Vosk, T., *Measurement Uncertainty*, in The Encyclopedia of Forensic Sciences, p.322-331 (2nd ed. Elsevier—2013)

- Vosk, T., *Metrological Epistemology*, in Understanding DUI Scientific Evidence, Ch. 7 & 8 (6th ed. Aspatore, 2013)

- Vosk, T., *Measurement Uncertainty: Forensic Metrology and Result Interpretation*, in Understanding DUI Scientific Evidence, Ch. 7 & 8 (4th ed. Aspatore, 2011)

- Vosk, T., *DUI Evidence and the National Academy of Sciences' Report on Forensic Science* in Understanding DUI Scientific Evidence, Ch. 5 (3rd ed. 2010).

- Vosk, T., *Uncertainty in Forensic Breath Alcohol Testing* in Intoxication Test Evidence, Ch. 56 (E. Fitzgeral ed., 2nd ed. 2009).

- National Academy of Sciences, *Strengthening Forensic Science in the United States: A Path Forward*, (2009).

VII. Papers & Articles

- Boscia, C., *Strengthening Forensic Alcohol Analysis in California DUI Cases: A Prosecutor's Perspective* 53 Santa Clara L. Rev. 733, 766 (2013).

- Christensen, A., et. al., Error and its Meaning in Forensic Science 59(1) J. FOR. SCI. 123 (2014).

- Ehrlich, C., Dybkaer, R., and Wöger, W., *Evolution of philosophy and description of measurement (preliminary rational for VIM 3)* 12 ACCRED. QUAL. ASSUR. 201 (2007).

- Estler, W.T., *Measurement as Inference: Fundamental Ideas* 48(2) ANNALS CIRP 611 (1999).

- Gullberg, G., *Estimating the Measurement Uncertainty in Forensic Blood Alcohol Analysis* 36 J. ANALYTICAL TOX. 153 (2012).

- Gullberg, G., *Estimating the Uncertainty Associated with Widmark's Equation as Commonly Applied in Forensic Toxicology*, 172 FOR. SCI. INT. 33 (2007).

- Gullberg, G., *Estimating the measurement uncertainty in forensic breath-alcohol analysis*, 11 ACCRED. QUAL. ASSUR. 562 (2006).

- Imwinkelried, E., *Forensic Metrology: The New Honesty about the Uncertainty of Measurements in Scientific Analysis*, UC Davis Legal Studies Research Paper Series, Research Paper No. 317 (Dec., 2012), available at http://papers.ssrn.com/sol3/papers.cfm?abstract_id=2186247.

- Jackson, et. al., *Backtracking Booze with Bayes-the Retrospective Interpretation of Blood Alcohol Data*, 31 BR. J. CLIN. PHARMAC. 55 (1991).

- Jones, A. W., *Dealing with Uncertainty in Chemical Measurements*, 14(1) NEWSL. OF THE INT. ASSOC. FOR CHEM. TEST. 6 (2003).

- Kristiansen, J. and Petersen, H.W., *An Uncertainty Budget for the Measurement of Ethanol in Blood by Headspace Gas Chromatography*, 28(6) J. ANAL. TOX. 456 (2004).

- Ku, H.H., *Notes on the Use of Propagation of Error Formulas*, 70(4) J. RES. NAT. BUR. OF STANDARDS 263 (1966).

- Posey and Mozayani, *The Estimation of BAC, Widmark Revisited*, 3 FOR. SCI., MED. & PATH. 33 (2007).

- Sklerov, J., & Couper, F., Calculation and Verification of Blood Ethanol Measurement Uncertainty for Headspace Gas Chromatography, 35(7) J. FOR. SCI. 402 (2011).

- Vosk, T., et al,. *The Measurand Problem in Breath Alcohol Testing*, __ J. FOR. SCI. __ (*In Press – May 2014*).

- Vosk, T., *Trial by Numbers: Uncertainty in the Quest for Truth and Justice*, THE NACDL CHAMPION, Nov. 2010.

- Vosk, T., *Chaos Reigning: Breath Testing and the Washington State Toxicology Lab*, The NACDL Champion, June 2008.

- Vosk, T., *Down the Rabbit Hole: The Arbitrary World of the Washington State Toxicology Lab* 22(2) Wash. Crim. Def. 37 (May 2008)

- Vranish, R., & Gullberg, R., Statistical Application of the "Bootstrap" Technique in Forensic Breath Alcohol Testing, 24(3) NEWSL. OF THE INT. ASSOC. FOR CHEM. TEST. 4 (2013).

- Wallace, J., *Ten Methods for Calculating the Uncertainty of Measurement*, 50(4) SCI. & JUSTICE 182 (2010).

- Wallace, J., *Proficiency Testing as a Basis for Estimating Uncertainty of Measurement Application to Forensic Alcohol and Toxicology Quantitations*, 55(3) J. FOR. SCI. 767 (2010).

VIII. Useful web based tools

- NIST Reference on Constants, Units and Uncertainty: http://physics.nist.gov/cuu/index.html

- NIST Traceability: http://ts.nist.gov/traceability/

- NIST Engineering Statistics Handbook: http://www.itl.nist.gov/div898/handbook/

- Elementary Concepts in Statistics: http://www.statsoft.com/textbook/stathome.html?stbasic.html&1

- Statistics and Science: Monograph Series: http://projecteuclid.org/DPubS?service=UI&version=1.0&verb=Display&handle=euclid.lnms/1215091126

- Web Pages that Perform Statistical Calculations: http://statpages.org/index.html

- Short Tandem Repeat DNA Internet DataBase: http://www.cstl.nist.gov/div831/strbase/
- DNA Advisory Board Quality Assurance Standards for Forensic DNA Testing Laboratories: http://www.cstl.nist.gov/strbase/dabqas.htm
- EURACHEM Guides and Documents: http://www.eurachem.org/guidesanddocuments.htm
- Forensic Science Resources on the Internet: http://www.istl.org/03-spring/internet.html
- DNA Forensic Mathematics: http://dna-view.com/
- Forensic Statistics and Legal Reasoning: http://www.josephbell.org/
- Quantifying Uncertainty in Analytical Measurement: http://www.measurementuncertainty.org/index.html.
- StatPages.org: http://statpages.org/index.html.
- SISA: http://www.quantitativeskills.com/sisa/index.htm.

IX. Journals – Free Access

- Journal of Research of NIST: http://nvl.nist.gov/nvl3.cfm?doc_id=89&s_id=117
- Measurement Science Review: http://www.measurement.sk/
- Metrology and Measurement Systems: http://www.metrology.pg.gda.pl/
- Pure and Applied Chemistry: http://www.iupac.org/publications/pac/index/
- The Journal of Philosophy, Science and the Law: http://www6.miami.edu/ethics/jpsl/index.html

APPENDIX C

TED VOSK - CV

CURRICULUM VITAE

Name: Ted Vosk

Address: 8105 NE 140th Pl.
Kirkland, WA 98034

Phone: (425) 753-6343
E-mail: tvosk@comcast.net

EDUCATION

University of Washington
- Course work in cybersecurity. 2006
- Graduate courses in physics. 2001-2003

Harvard Law School 1999
Juris Doctorate
- Thesis: Human Cloning and FDA Regulation.

Cornell University 1995-1996
- Graduate studies in Physics, Doctorate program.

Eastern Michigan University 1995
Bachelor of Science – Magna cum laude, Physics and Mathematics w/ minor in Astronomy.
- Honors in Theoretical Physics.
- Honors in Mathematics.
- Honors from University Honors College.
  - Honors Thesis: A Comparative Spectroscopic and Topographic Analysis of the Surface of Graphite.
- Phi Kappa Phi: Academic Honors Society.
- Sigma Pi Sigma: Physics Honors Society.
- Golden Key National Honor Society.
- The Stoic Society: Academic Honors Society.

University of Michigan, Space Physics Research Lab 1993
- Research Associate: Modeling atomic oxygen concentration and dynamics in the mesosphere and thermosphere.

SELECTED FORENSICS TRAINING

Advanced DNA Mixture Interpretation and Statistical Approaches (2-day workshop). American Academy of Forensic Sciences. Atlanta, GA, Feb., 2012.

Bayesian Networks in Forensic Science (1-day course). 8th International Conference on Forensic Inference and Statistics. Univ. of Wash. Dept. of Biostatistics and Schools of Public Health and Law. Seattle, WA, July, 2011.

Introduction to Uncertainty in Forensic Chemistry and Toxicology I, II, III (Online training). National Institute of Justice/RTI International. Oct. 2010.

Good Measurement Practices in the Proper Use and Calibration of Balances and Pipettes (1-day workshop). American Academy of Forensic Sciences. Seattle, WA, Feb., 2010.

Selected Forensics Training (cont.)

ISO/IEC 17025:2005 Section 5.4.6: Estimation of Uncertainty – Is Anyone Certain What This Means? (1-day workshop). American Academy of Forensic Sciences. Denver, CO, Feb., 2009.

INTOXILYZER 8000 Operator's Course (3-day course). Faculty: Dr. Alfred Staubus, College of Pharmacy (Pharmaceutics), Ohio State University. New Orleans, LA, Sept., 2006.

NHTSA/IACP Standardized Field Sobriety Testing Instructor Course (4-day course). Walden, Platt & Associates. San Antonio, TX, June, 2005.

NHTSA Standardized Field Sobriety Testing Practitioner Course (3-day course). Walden, Platt & Associates. Seattle, WA, May, 2005.

Drug Evaluation and Classification (DECP) Overview Course (3-day course). Walden, Platt & Associates. Seattle, WA, Oct., 2004.

Law

Attorney/Consultant/Writer, Seattle, WA, 2004-Present.
Criminal Defense/Appeals/Forensic metrology.
Of Counsel, Cowan, Kirk, Gaston (2009-present).
Of Counsel, Callahan Law (2007-2009).

Attorney, Magnuson Lowell, Redmond, WA, 2003-2004.
Criminal Defense/Tort/Administrative Law.

Public Defender, Tucker & Stein, Bellevue, WA, 2001.

Deputy Prosecutor, City of Redmond, Redmond, WA, 2000-2001.

Science Research Management

Univ. of Washington, Dept. of Chemistry, Seattle, WA, 2001-2003.
- Acting Managing Director, 2002-2003.
  National Science Foundation, Science & Technology Center on Materials and Devices for Information Technology Research.
- Research Program Manager, 2001-2003.
  Dalton Research Group.

Teaching

Contributing Faculty, 2011-2012.
On Demand Science Education Project, Jan Semenoff International.

Guest Lecturer, 2009.
Univ. of Washington, Evans School of Public Affairs, Seattle, WA.
Topics in Science, Technology, and Public Policy: Policy Formulation and Implementation.

Guest Lecturer, 2007-2008.
Edmunds Community College, Edmunds, WA.
BUS 240: Business Law.

Science Outreach/Nonprofit

Vice President, Writer, Broadcaster, Public Speaker, 2003-2008.
Celestial North, Seattle, WA.
Educating K-12 and lay public in astronomy and space science.
"It's Over Your Head" radio program broadcast on KSER 90.7 FM.

Boards

Faculty Advisory Board, 2011-present.
BAC Tracker International, Inc.

Legislative Activities

Forensic Investigations Council Legislation Workgroup, 2009-2010.
Washington Association of Criminal Defense Lawyers
Workgroup drafted forensic reform legislation passed into law.

Admitted to Legal Practice

United States Supreme Court, 2007.
United States Court of Appeals for the Ninth Circuit, 2007.
United States District Court, Western District of Washington, 2003.
State of Washington, 2000.
State of Massachusetts, 1999.
State of California (Pro Hac Vice – Trial counsel), 2011.
State of Oregon (Pro Hac Vice – Trial and Appellate counsel), 2006-2008.

Notable Cases

1. State v. Carson, No. 12-01408 SD (55th Dist. Ct. Ingham Co. MI – 1/8/14) (counsel Michael Nichols; consultant Ted Vosk).
2. State v. King County Dist. Court West Div., 307 P.3d 765 (Wash.App. 2013).
3. State v. Olson, 170 Wash.App. 1032 (Wash.App. 2012).
4. *People v. Jabrocki*, No. 08-5461-FD (79th Dist. Ct. Mason Co. MI – 5/6/11) (counsel Michael Nichols; consultant Ted Vosk)
5. Washington State Forensic Uncertainty Litigation:
   State v. Fausto, No. C076949 (King Co. Dist. Ct. WA – 9/20/10).
   State v. Weimer, No. 7036A-09D (Sno. Co. Dist. Ct. WA – 3/23/10).
6. Seattle v. Winebrenner, 219 P.3d 686 (Wash. 2009): Amicus Curiae – WACDL.
7. Washington State Toxicology Lab Litigation:
   State v. Ahmach, # C00627921 (King Co. Dist. Ct. – 1/30/08).
8. Ludvigsen v. City of Seattle, 174 P.3d 43 (Wash. 2007).
9. City of Fircrest v. Jensen, 143 P.3d 776 (Wash. 2006).
10. Devine v. Dept. of Licensing, 110 P.3d 237 (Wash. App. 2005): Consultant.
11. City of Seattle v. Clark-Munoz, 93 P.3d 141 (Wash. 2004): Consultant.

Professional Memberships

American Academy of Forensic Sciences (AAFS), Fellow
American Association for the Advancement of Science (AAAS)
American Physical Society (APS)
- Topical Group on Precision Measurements and Fundamental Constants
- Topical Group on Instrument and Measurement Science

Professional Memberships (cont.)

American Mathematical Society (AMS)
American Chemical Society (ACS)
International Union of Pure and Applied Chemistry (IUPAC)
American Society for Testing and Materials International (ASTM)
- Workgroup for the Revision of ASTM standard E620, 2011-present.

Mensa
National Association of Criminal Defense Lawyers (NACDL)
National College for DUI Defense (NCDD)
Washington Association of Criminal Defense Lawyers (WACDL)
- Legislative Committee Member, 2006-2010
- Forensic Investigations Council Legislation Workgroup, 2009-2010

Honors & Awards

1. Super Lawyer, Washington Law & Politics Magazine, 2012, 2013.
2. Top Attorneys in Washington, Seattle Metropolitan Magazine, 2012, 2013.
3. VIP Top Fundraiser – Race for a Cure, Crohn's & Colitis Foundation of America, 2012.
4. Outstanding NCDD Member, National College for DUI Defense, 2011.
5. Recognition for contributions & instruction, National College for DUI Defense, 2011.
6. President's Award, Wash. Assoc. of Criminal Defense Lawyers, 2008.
7. Pro Bono Public Service Commendation, Washington State Bar Association, 2006, 2007, 2008.
8. Certificate of Distinction, Washington Foundation for Criminal Justice, 2007.
9. Out of This World Award for Excellence in Astronomy Outreach, Astronomy Magazine, 2006.
10. Super Lawyer Rising Star, Washington Law & Politics Magazine, 2005.
11. Goldwater Scholar in Mathematics, Science and Engineering, Goldwater Foundation, 1993-1995.
12. National Deans List, 1995.
13. Outstanding Student of Mathematics, E. Mich. Univ. Dept. of Mathematics, 1995.
14. Harcourt Brace Book Award: Outstanding Scholarship in Physics, E. Mich. Univ. Dept. of Physics, 1995.
15. Campus Leader Scholarship, E. Mich. Univ., 1994.
16. Leib Scholarship, E. Mich. Univ. Dept. of Physics, 1992, 1993, 1994.
17. Lobbestael Scholarship, E. Mich. Univ. Dept. of Mathematics, 1993, 1994.
18. Recognition of Excellence Scholarship, E. Mich. Univ. Col. Of Arts & Sciences, 1992, 1993, 1994.
19. Robert Silver Award, E. Mich. Univ. Dept. of Physics, 1993.

Publications - Book Chapters/Texts

1. Forensic Metrology: Scientific Measurement and Inference for Lawyers, Judges, and Criminalists (CRC/Taylor Francis Group – *Available September 2014*).
2. Measurement Uncertainty, Encyclopedia of Forensic Sciences, p.322-331 (2nd ed. Elsevier – 2013).

PUBLICATIONS - BOOK CHAPTERS/TEXTS (CONT.)

3. Metrological Epistemology, Understanding DUI Scientific Evidence, (6th ed. Aspatore, 2013).
4. An Introduction to Uncertainty, Mississippi DUI: Law & Practice, Ch. __ (Thomson-West – 2013).
5. The Science of Measurement Uncertainty, Defending Drinking Drivers, § 210 (James Publishing, 2012).
   - §210.2: Uncertainty in Bodily Alcohol Chemical Tests.
   - §210.3: Uncertainty in Breath Tests.
   - §210.4: Uncertainty in Blood Tests.
   - §210.4: Application of Measurement Uncertainty to the Admissibility of Chemical Tests.
6. Forensic Metrology, Drunk Driving Defense, § 6.04K (7th ed. Aspen, 2012).
7. Computational Aspects of Measurement Uncertainty in Washington State Breath Alcohol Tests, Washington DUI Practice Manual, ch. 25A (Wash Prac. Series, v.32)(Thomson-West, 2011 & 2012 Update).
8. Measurement Uncertainty: Forensic Metrology and Result Interpretation, ch. 7 & 8, Understanding DUI Scientific Evidence, (4th ed. Aspatore, 2011).
   - Chapter 7: Part I, Measurement Results and Interpretation.
   - Chapter 8: Part II, Legal Analysis.
9. DUI Evidence and the National Academy of Sciences' Report on Forensic Science, Understanding DUI Scientific Evidence, ch. 5, app. E, F & G (3rd ed. Aspatore, 2010).
10. Scientific Principles of Breath Alcohol Testing, Defending DUI's in Washington, ch. 13 (3rd ed. LexisNexis 2010).
11. The DataMaster, Defending DUI's in Washington, ch. 13/13A (3rd ed. LexisNexis 2008, 2009 & 2010 Update).
12. Direct Examination of the Defense Expert, Defending DUI's in Washington, ch. 15 (3rd ed. LexisNexis 2008, 2009 & 2010 Update).
13. Uncertainty in Forensic Breath Alcohol Testing, Intoxication Test Evidence, ch. 56 (2nd ed. Thomson-West, 2009).
14. Field Sobriety Testing and Driver Impairment: Linked or Not?, Understanding DUI Scientific Evidence, ch. 3, app. A & B (2nd ed. Aspatore, 2009).
15. Standardized Field Sobriety Testing, Washington DUI Practice Manual, ch. 21 (Wash Prac. Series, v.32)(Thomson-West, 2008).
16. Breath test/toxicology lab argument summary, Washington DUI Practice Manual, § 32-11, (Wash Prac. Series, v.32)(Thomson-West, 2008).

PUBLICATIONS - ARTICLES

1. The Measurand Problem in Breath Alcohol Testing, Journal of Forensic Science, (*In Press* – May 2014).
2. Errors and Uncertainties: What Hath the GUM Wrought?, Proceedings of ASME 2013 International Mechanical Engineering Congress, *In Press*.
3. Uncertainty in the Quest for Justice, Wash. Crim. Def., Nov. 2013.
4. Trial by Numbers: Uncertainty in the Quest for Truth and Justice, The NACDL Champion, Nov. 2010 (Reprinted in The Voice for the Defense, April 2011).
5. Forensic Evidence Checklist: Steps to Take to Protect Your Client From Bad Science, Wash. Crim. Def., Nov. 2010.

PUBLICATIONS – ARTICLES (CONT.)

6. Chaos Reigning: Breath Testing and the Washington State Toxicology Lab, The NACDL Champion, June 2008.
7. Down the Rabbit Hole: The Arbitrary World of the Washington State Toxicology Lab, Wash. Crim. Def., May 2008.
8. Due Process and Science by Legislative Decree, Wash. Crim. Def., Feb. 2007.
9. Field Sobriety Tests: Another Government Lie?, (contract article) Bar News, August 2007.
10. A New Paradigm for Challenging Breath Test Evidence in Washington, (contract article) Bar News, June 2007.
11. Precluding Standardized Field Sobriety Tests in non-per se Prosecutions, Wash. Crim. Def., Feb. 2006.

ABSTRACTS:

1. The Measurand Problem in Breath Alcohol Testing. Proceedings of the American Academy of Forensic Sciences 66th Annual Scientific Meeting, p.273, 2014.
2. The Measurand Problem in Infrared Breath Alcohol Testing. Bulletin of the American Physical Society, APS March Meeting 57(1), 2012.
3. Uncertainty Analysis in Forensic Practice: How to Apply It Wherever Scientific Integrity Demands Its Use. Proceedings of the American Academy of Forensic Sciences 64th Annual Scientific Meeting, p.__, 2012.
4. Physics in the Courtroom (invited paper), Bulletin of the American Physical Society, 13th Annual Meeting of the Northwest Section of the APS 56(10), 2011.

WORKSHOPS:

1. Science, Law, and the Inferential Process: The Epistemology of Scientific Conclusions and Their Role in the Legal Fact-Finding Process. Chair: Ted Vosk. Co-Chair: Hon. Rod Kennedy. 66th Annual Scientific Meeting of the American Academy of Forensic Sciences. Seattle, WA. Feb., 2014.
2. Forensic Metrology - The Science and Practicalities of Forensic Measurements in the Courtroom. Washington Defender Association/King County Dept. of Public Defense. Seattle, WA. Jan. 2014.
3. Courtroom Challenges for Forensic Scientists. 35th Annual Conference, Southwestern Association of Forensic Scientists. Santa Fe, NM. Oct., 2013.
4. Science in the Courtroom: A Matter of Perspective? Chair: Ted Vosk. Co-Chair: Hon. Rod Kennedy. 65th Annual Scientific Meeting of the American Academy of Forensic Sciences. Washington, D.C., Feb., 2013.
5. Blood Alcohol Concentration (BAC) Evidence: Extrapolation, Interpretation, and Testimony in the Post–NAS Era. Co-Chairs: Hon. Rod Kennedy & A.W.R. Forrest. 63rd Annual Scientific Meeting of the American Academy of Forensic Sciences. Chicago, Ill. Feb., 2011.
6. Attorneys and Scientists in the Courtroom: Bridging the Gap. Co-Chairs: Ted Vosk & Max Houck. 62nd Annual Scientific Meeting of the American Academy of Forensic Sciences. Seattle, WA. Feb., 2010.

PRESENTATIONS

1. Scientific Method as Applied Epistemology. 66th Annual Scientific Meeting of the American Academy of Forensic Sciences. Seattle, WA. Feb., 2014.
2. Decision Theory: A Quantitative Comparison of Scientific and Legal Inference. 66th Annual Scientific Meeting of the American Academy of Forensic Sciences. Seattle, WA. Feb., 2014.
3. The Measurand Problem in Breath Alcohol Testing. 66th Annual Scientific Meeting of the American Academy of Forensic Sciences. Seattle, WA. Feb., 2014.
4. Epistemological Structure of Scientific Conclusions: Finding Out What Scientific Results Really Mean. Forensic Metrology - The Science and Practicalities of Forensic Measurements in the Courtroom. Washington Defender Association/King County Dept. of Public Defense. Seattle, WA. Jan. 2014.
5. Measurement Uncertainty. Forensic Metrology - The Science and Practicalities of Forensic Measurements in the Courtroom. Washington Defender Association/King County Dept. of Public Defense. Seattle, WA. Jan. 2014.
6. An Introduction to Measurement Uncertainty in Chemical Measurement. Blurred Lines: Advanced DWI Defense. New Mexico Criminal Defense Lawyers Association. Albuquerque, NM. Oct. 2013.
7. Litigating the Issue of Measurement Uncertainty in Chemical Test Cases. Blurred Lines: Advanced DWI Defense. New Mexico Criminal Defense Lawyers Association. Albuquerque, NM. Oct. 2013.
8. Uncertainty in Breath and Blood Alcohol Tests. Pierce County Office of Public Defense. Tacoma, WA. Oct. 2013.
9. Metrological Epistemology: Science and Matters of Belief. DUI Boot Camp, "Gamble On Detroit." Michigan Association of OWI Attorneys. Detroit, MI. Sept. 2013.
10. Measurement uncertainty and DNA litigation. Science in the Courtroom for the 21st Century: Current Issues in DNA Litigation. Chicago, IL. March 2013.
11. Introduction to Forensic Metrology. 7th Annual - Diving Into An OVI / DUI Case. Key West, FL. Feb. 2013.
12. Science: Information to Inference. AAFS Annual Meeting, American Academy of Forensic Sciences. Washington, D.C., Feb., 2013.
13. Science in the Courtroom: A Defense Perspective. 65th Annual Scientific Meeting of the American Academy of Forensic Sciences. Washington, D.C., Feb., 2013.
14. Metrology for Lawyers. Occupy DUI. California DUI Lawyers Association. San Diego, CA. Nov. 2012.
15. Uncertainty, Error & Inference in the Courtroom. Forensics: From Crime Scene to Courtroom. North Carolina Advocates for Justice. Raleigh, NC. April, 2012.
16. The Measurand Problem in Infrared Breath Alcohol Testing. American Physical Society Annual Meeting, Boston, MA. Feb., 2012.
17. Uncertainty Analysis in Forensic Practice: How to Apply It Wherever Scientific Integrity Demands Its Use. AAFS Annual Meeting, American Academy of Forensic Sciences. Atlanta, GA. Feb., 2012.

PRESENTATIONS (CONT.)

18. Draeger Schmaeger…It's Still Just A Machine (Introduction to Infrared and Fuel Cell Based Breath Testing). Defending DUIs, Washington Foundation for Criminal Justice. SeaTac, WA. Dec., 2011.
19. Physics in the Courtroom (invited paper). American Physical Society – Northwest Section Annual Meeting, Corvallis, OR. Oct. 2011.
20. Introduction to Uncertainty, National College for DUI Defense. Webinar, Sept. 2011.
21. Measurement Uncertainty: Meaning and Application. Summer Session, National College for DUI Defense. Cambridge, MA, July 2011.
22. Scientific Measurements/Uncertain Justice: What the Numbers Actually Mean. Advanced Skills for the Experienced Practitioner. Supreme Court of Virginia, The Chief Justice's Indigent Defense Training Initiative. Richmond, VA, April 2011.
23. State v. Fausto - Trial by Numbers: Bias & Uncertainty in the Quest for Justice. Making Sense of Science. National Association of Criminal Defense Lawyers. Las Vegas, NV, March, 2011.
24. Tiny Science: Principles and Foundations for Lawyers in the Courtroom. DUI Strategies for Defense. Washington Defender Association. Seattle, WA. March, 2011.
25. Limitations and Uncertainty Management in Expert Interpretation and Testimony of Extrapolated BAC Results. 63rd Annual Scientific Meeting of the American Academy of Forensic Sciences. Chicago, Ill. Feb., 2011.
26. Methods for Interpreting and Extrapolating BAC Results using Computerized Models. 63rd Annual Scientific Meeting of the American Academy of Forensic Sciences. Chicago, Ill. Feb., 2011.
27. Fryebert – The Standards for Admissibility of Scientific Evidence in Washington State. Challenging Bad Science – Putting the NAS Report to Work, Washington Association of Criminal Defense Lawyers. Seattle, WA, Oct., 2010.
28. Error and Uncertainty in the Quest for Justice. Challenging Bad Science – Putting the NAS Report to Work, Washington Association of Criminal Defense Lawyers. Seattle, WA, Oct., 2010.
29. Trial by Numbers: Uncertainty in the Quest for Justice. DataMaster and Scientific Evidence in DWI/DUI Cases, South Carolina Association of Criminal Defense Lawyers. Myrtle Beach, SC, Oct., 2010.
30. Confidence Intervals to Interpret and Challenge Breath Test Results. Washington Defender Association/Northwest Defender Association. Seattle, WA, July, 2010.
31. Forensic Science in the Courtroom: Uncertainty, Error and Validity. 7th Annual DEA Southwest Laboratory Symposium, U.S. Drug Enforcement Agency Southwest Laboratory. Vista, CA, May, 2010.
32. Forensic Metrology - How to Challenge the Scientifically Unacceptable Naked Test Result. 23rd Annual Aggressive Defense of the Accused Impaired Driver, Arizona Attorneys for Criminal Justice. Tucson, AZ, May, 2010.
33. Metrology for Lawyers: Why the Numbers Matter and How to Give Them Life. Mastering Scientific Evidence, Texas Crim. Def. Lawyers Assoc. & Nat'l College for DUI Defense. New Orleans, LA. April, 2010.

PRESENTATIONS (CONT.)

34. Limits of Methodology: Calibrations, Confidence Limits, and Measuring Accuracy. Skeptically Determining the Limits of Expert Testimony and Evidence: An Investigation of Scope, Expertise and Process, State Bar of New Mexico. Albuquerque, NM. March, 2010.
35. Forensic Metrology - The Foundation of Science in the Courtroom. "Fast Eddie" Kuwatch All Day DUI Seminar, California DUI Lawyers Association. San Francisco, CA. March, 2010.
36. Metrology: A Knowledge Base for Communication and Understanding. 62nd Annual Scientific Meeting of the American Academy of Forensic Sciences. Seattle, WA. Feb., 2010.
37. Fun With Bodily Fluids: How to Make the Toxicology Lab Your Best Witness. Discovery: It's Elementary, Washington Association of Criminal Defense Lawyers. Bellevue, WA. Feb., 2010.
38. Breath Testing – Beating the Odds. Defending DUIs, Washington Foundation for Criminal Justice. SeaTac, WA. Dec., 2009.
39. Forensic Metrology: A 1-day course. Defending DUIs, Spokane County Public Defenders Office. Spokane, WA. Sept., 2009.
40. Forensic Metrology: The Basics. Washington Defender Association. Seattle, WA. Sept., 2009.
41. Forensic Metrology – Why it's Essential to Litigating Forensic Science Cases. Science in the Courtroom for the 21st Century: Issues in Forensic DNA, DePaul College of Law, Center for Law and Science. Chicago, Ill. May, 2009.
42. Forensic Metrology: The Key to the Kingdom. National Forensic Blood and Urine Testing Seminar, Georgia Association of Criminal Defense Lawyers. San Diego, CA. May, 2009.
43. The National Academy of Science's Report and the Future of Forensic Science in Washington. Cars and Crime: Electrifying Your Defense, Washington Association of Criminal Defense Lawyers. Seattle, WA. March, 2009.
44. Are Breath Test Results Accurate? Defending DUIs, Snohomish County Bar Association. Everett, WA. Nov., 2008.
45. Washington State Toxicology Lab Controversy. Annual Conference, Washington Court Reporters. Tacoma, WA. Oct., 2008.
46. Toxicology Lab Results — Where do we go from here? The 15th Annual Criminal Justice Institute, Washington State Bar Association. Seattle, WA. Sept., 2008.
47. DUI/Crime Lab Update. WACDL Annual Conference: Equal Justice for ~~All~~ Some?, Washington Association of Criminal Defense Lawyers. Chelan, WA. June, 2008.
48. Scientific Standards – How to Master Science Without Being a Scientist. Northwest Defenders Association. Seattle, WA. June, 2008.
49. Motions – Crippling the State's Case. Defending DUIs, Washington Foundation for Criminal Justice. SeaTac, WA. Dec., 2007.
50. Challenging the Crime Lab. District Court Practice, Washington Association of Criminal Defense Lawyers. Tacoma, WA. Oct., 2007.
51. The State Toxicology Lab. Basics of DUI Defense: From the Lab to the Courtroom, Washington Defender Association. Seattle, WA. Oct., 2007.

PRESENTATIONS (CONT.)

52. Breath Tests and the State Toxicology Lab. State Toxicology Lab: What Defenders Need to Know, Washington Defender Association. Seattle, WA. Sept., 2007.
53. The Washington State Toxicology Lab – What you need to know. DUIs: From Traffic Stop to Trial, Spokane County Bar Association. Seattle, WA. Sept., 2007.
54. Techniques, Strategies and Methods for Addressing Breath Tests. DUIs In a Post-Jensen World, Washington Defender Association. Seattle, WA. Feb., 2007.
55. The Mathematical Analysis of Breath Testing. DUIs for Public Defenders, Washington State Office of Public defense. Bellevue, WA. March, 2006.
56. Standardized Field Sobriety Testing. DUIs for Public Defenders, Washington State Office of Public defense. Bellevue, WA. March, 2006.
57. Strings, Gravity, and Locality: An Overview of Modern Cosmology. Table Mountain Star Party, Table Mountain Star Party Association, Ltd. Ellensburg, WA. August, 2005.
58. Constitutional Separation of Powers. A Constitutional Challenge to the 2004 DUI Law, Washington Alliance of DUI Lawyers. Seattle, WA. Dec., 2004.
59. Celestial North: Awe, Wonder and the Need to Know. Table Mountain Star Party, Table Mountain Star Party Association, Ltd. Ellensburg, WA. August, 2004.
60. Site Asymmetry in Scanning Tunneling Microscopy of Graphite. Undergraduate Research Symposium XV, Eastern Michigan University College of Arts and Sciences. Ypsilanti, MI. June, 1995.
61. Site Asymmetry in Scanning Tunneling Microscopy of Graphite. National Undergraduate Research Symposium, Argonne National Laboratory. Argonne, IL. Nov., 1994.
62. Communication via Laser Beam. Undergraduate Research Symposium XIII, Eastern Michigan University College of Arts and Sciences. Ypsilanti, MI. June, 1993.
63. Photoelectric Photometry of Small Amplitude Red Variables. Undergraduate Research Symposium XII, Eastern Michigan University College of Arts and Sciences. Ypsilanti, MI. June, 1992.

OTHER SPEAKING/EVENT ACTIVITIES

1. Panelist: Discussion and perspectives on statistical issues. Science in the Courtroom for the 21st Century: Current Issues in DNA Litigation. Chicago, IL. March 2013.
2. Graduation Speaker: Snohomish County Youth Drug Court Graduation. Snohomish, WA. April, 2012.
3. Panelist: Preparing for Expert Testimony at Trial: A Professionalism Discussion. Skeptically Determining the Limits of Expert Testimony and Evidence: An Investigation of Scope, Expertise and Process, State Bar of New Mexico. Albuquerque, NM. March, 2010.
4. Session Moderator: DNA II. AAFS Annual Meeting, American Academy of Forensic Sciences. Seattle, WA. Feb., 2010.

OTHER SPEAKING/EVENT ACTIVITIES (CONT.)

5. Panelist: The Implications and Opportunities of the National Academy of Sciences Report on Forensic Science. Science in the Courtroom for the 21st Century: Issues in Forensic DNA, DePaul College of Law, Center for Law and Science. Chicago, Ill. May, 2009.
6. Panelist: Decisions From Around the USA to Help You Win Tough Cases. National Forensic Blood and Urine Testing Seminar, Georgia Association of Criminal Defense Lawyers. San Diego, CA. May, 2009.
7. Judge, 1L Mock Trial Competition. University of Washington, School of Law. Seattle, WA. Oct., 2008.
8. Organizer. Basics of DUI Defense: From the Lab to the Courtroom. Spokane, WA. Oct. 19, 2007.
9. Organizer. Basics of DUI Defense: From the Lab to the Courtroom. Seattle, WA. Oct. 12, 2007.
10. Co-author: Field Sobriety Tests: "Tell Them No!" ATLA Annual Convention, DUI Defense Tips From the Masters, Association of Trial Lawyers of America. Seattle, WA. July, 2006.
11. Co-author/Editor: The Bad-Bad Case and Blood Tests. Presented at National College for DUI Defense, Summer Session, National College for DUI Defense. Cambridge, MA. July, 2005.
12. Moderator: Physics and Astronomy. Undergraduate Research Symposium XXV, Eastern Michigan University College of Arts and Sciences. Ypsilanti, MI. April, 2005.
13. Emcee: Undergraduate Research Symposium XV, Eastern Michigan University College of Arts and Sciences. Ypsilanti, MI. June, 1995.

TELEVISION/RADIO/PRINT FEATURES:

1. Interview: Consequences of the King County District Court uncertainty decision.
   Schram On The Story: FM 97.7 KOMO (9/23/10).
2. Interview: If we can't use breathalyzers, how do we keep drunks off the road? The Dave Ross Show: AM 710 KIRO (8/6/08).
3. Interview: Problems at the Washington State Toxicology Lab.
   Up Front with Robert Mak: KING 5 (2/2/08).
4. Profile: Harvard Law Bulletin, *Celestial Reasonings*, Spring 2007, at 52.
5. Profile: The Ann Arbor News, *EMU graduate finishes like Rocky after rocky start*, April 23, 1995, at A1.