Forensic Metrology

A Primer for Lawyers, Judges and Forensic Scientists

Ted Vosk

Forensic Metrology
A Primer for Lawyers, Judges and Forensic Scientists

Based on the textbook:
Vosk, Emery, Fitzgerald, *Forensic Metrology: A Primer on Scientific Measurement for Lawyers, Judges and Forensic Scientists* (CRC Press – In Preparation)

Ted Vosk

v 1.6

8105 NE 140th Pl., Bothell WA 98011

For Kris

8105 NE 140th Pl., Bothell WA 98011

Table of Contents

WHY FORENSIC METROLOGY FOR LAWYERS, JUDGES AND FORENSIC SCIENTISTS? i
I. CONCERNING FORENSIC SCIENCE AND THE ADMINISTRATION OF JUSTICE 1
A. SCIENCE 1
B. LAWYERS AND JUDGES 2
II. WHERE TO BEGIN? 2
A. SCIENCE 101 2
B. METROLOGY 3
III. METROLOGY: A LITTLE BACKGROUND 3
A. METROLOGIST 3
B. BRIEF HISTORY 4
IV. METROLOGY 4
A. METROLOGICAL FOCUS 4
B. WEIGHTS & MEASURES 5
C. MEASUREMENT AND TESTING PROCESS 6
D. QUALITY ASSURANCE 9
E. MEASUREMENT INTERPRETATION 14
F. SCIENTIFIC STANDARDS 35
V. FORENSIC METROLOGY 37
A. FORENSIC METROLOGY 37
B. FORENSIC WEIGHTS AND MEASURES 37
C. FORENSIC MEASUREMENT AND TESTING PROCESS 39
D. FORENSIC QUALITY ASSURANCE 40
E. FORENSIC MEASUREMENT/OBSERVATION INTERPRETATION 43
F. SCIENTIFIC FORENSIC STANDARDS 52
VI. SPECIAL TOPICS 53
A. DNA 53
B. BLOOD ALCOHOL TESTING 55
VII. FORENSIC METROLOGY AND THE LAW 63
A. GENERAL PRINCIPLES 63
B. DAUBERT, FRYE AND EVID. R. 702 63
C. STATUTORY/REGULATORY METROLOGICAL PROVISIONS 78
VIII. APPLICATIONS 81
A. EVID. R. 702. – BLOOD TESTING: REPORTING RESULTS 81
B. STATUTORY/REGULATORY REQUIREMENTS – BLOOD TESTING: STANDARDS AS EVIDENCE 82
IX. CASE STUDY: WASHINGTON STATE TOXICOLOGY LAB – *State v Ahmach* 83
A. VIOLATED STANDARDS – EXAMPLES 83
B. EVIDENTIARY RULING 85
X. RECENT DECISION INCORPORATING NAS REPORT – *State v Berry* 86
XI. EXERCISES 88
APPENDIX A: MISCELLANEOUS RESOURCES ..... A

8105 NE 140th Pl., Bothell WA 98011

APPENDIX B: DECISION – *State v Ahmach*, King County District Court (Jan. 2008) .... B
APPENDIX C: DECISION – *State v Berry*, Pierce County District Court (Dec. 2009) C
APPENDIX D: TED VOSK - CV .... D

8105 NE 140th Pl., Bothell WA 98011

WHY FORENSIC METROLOGY FOR LAWYERS, JUDGES AND FORENSIC SCIENTISTS?[1]

"The ultimate mission of the system upon which we rely to protect the liberty of the accused as well as the welfare of society is to ascertain the factual truth."[2] "Complete, competent, and impartial forensic-science investigations can be that 'touchstone of truth' in a judicial process that works to see that the guilty are punished and the innocent are exonerated."[3]

Unfortunately, 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, inadequate training and a lack of resources have led to systemic failures to adhere to basic scientific standards, principles and practices. The situation has grown so bad that the National Academy of Sciences Report concludes that "[t]he law's greatest dilemma in its heavy reliance on forensic evidence…concerns the question of whether—and to what extent—there is *science* in any given 'forensic science' discipline."[4] Given the significant role scientific knowledge and evidence plays in the courtroom, this weakness threatens to undermine the integrity of our system of justice as a whole. It is "clear that change and advancements, both systemic and scientific, are needed in a number of forensic science disciplines—to ensure the reliability of the disciplines, establish enforceable standards, and promote best practices and their consistent application."[5]

Forensic science professionals 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

[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] *Commonwealth of Northern Mariana Islands v. Bowie*, 243 F.3d 1109, 1114 (9th Cir. 2001).
[3] Peterson, THE EVOLUTION OF FORENSIC SCIENCE: PROGRESS AMID THE PITFALLS 36 Stetson Law Rev. 621, 660 (2007).
[4] NATIONAL ACADEMY OF SCIENCES, *Strengthening Forensic Science in the United States: A Path Forward* [hereinafter NAS 2009], 87 (2009).
[5] NAS p., xix.

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.[6] Thus, if the integrity of our justice system is to be preserved, it is equally important to reform the practices of lawyers and judges. In today's technologically advanced society, the law's truth finding function cannot be achieved if its practitioners are ignorant of the basic tenants of science.

Exacerbating the situation is the fact that forensic scientists and legal professionals often seem to be peoples divided by a common language. For example, courts commonly treat the concept of accuracy as a binary question with a simple yes or no answer within the scientific community. Either the evidence or method is accurate or it is not. Scientists, on the other hand, often find this an untenable oversimplification that clouds the true meaning of a piece of evidence. To them, accuracy is a thing that can be objectively quantified but usually only subjectively adjudged. When these cultures collide in the courtroom, legal professionals often refer to scientist witnesses as deceptive or evasive while scientists may refer to legal professionals as dishonest, manipulative or unsophisticated. And again the integrity of the system suffers as both sides believe the other to be attempting to subvert the truth. If forensics is to live up to its promise in the courtroom, lawyers, judges and scientists must also be able to communicate in a common language.

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

[6] 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).

activities, the only thing that binds our notions of the physical world to reality is faith.[7] 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 would provide a discrete tool to guide the performance of certain aspects of *all* scientific procedures and the critical evaluation of certain aspects of *all* scientific claims as well as a common language for communication.

Metrology,[8] the science of measurement and observation, provides such principles and tools. Metrological[9] principles apply to every measurement and observation made in every lab anywhere on the

[7] 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.
[8] Including the emerging field of "proto-metrology".

planet. As noted physicist Lord Kelvin said over a century ago, "…if science is measurement, then without metrology there can be no science."[10] Thus, given a basic understanding of metrology, a scientist can properly employ the fundamental tools of science, even a nonscientist can begin to engage in a critical analysis of scientific claims across a broad spectrum based on metrological principles and the two groups can clearly communicate about these matters. Metrology, then, offers an operational, analytical and language knowledge base to facilitate forensic sciences ability to be a 'touchstone of truth' within the criminal justice system.

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, judges and forensic scientists? By gaining a basic understanding of metrological principles, forensic scientists better learn how to conduct science; judges and lawyers gain a basic understanding of science itself; and the language of metrology can provide a common vocabulary for communication. In the courtroom, all are empowered to be full participants in the presentation and critical analysis of forensic evidence. 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.

---

[9] Including the emerging field of "proto-metrology".
[10] 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. FORENSIC SCIENCE.

1. "Appropriate scientific standards are widely ignored in forensic laboratories"[11] contributing "to questions about the validity of conclusions."[12]

2. "Few forensic science methods have developed adequate measures of the accuracy of inferences made by forensic scientists."[13]

3. "Much forensic evidence…is introduced in criminal trials without any meaningful scientific validation, determination of error rates, or reliability testing."[14]

4. "[F]orensic scientists themselves often fail to consider or appreciate measurement uncertainty."[15]

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."[16]

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."[17]

7. "[B]ad laboratory practices have bedeviled even the FBI laboratories."[18]

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!"[19]

9. "The forensic science system, encompassing both research and practice, has serious problems"[20] making "clear that change and advancements, both systemic and scientific, are

---

[11] Erica Beecher-Monas, *Evaluating Scientific Evidence: An Interdisciplinary Framework for Intellectual Due Process*, 97 (Cambridge Press 2007).
[12] NAS, *Strengthening Forensic Science in the United States: A Path Forward*, 7-7 (2009).
[13] NAS, *Strengthening Forensic Science in the United States: A Path Forward*, 6-1 (2009).
[14] NAS, *Strengthening Forensic Science in the United States: A Path Forward*, 3-18 (2009).
[15] Gullberg, *Estimating the measurement uncertainty in forensic breath-alcohol analysis*, 11 ACCRED. QUAL. ASSUR. 562, 563 (2006).
[16] 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).
[17] NAS, *Strengthening Forensic Science in the United States: A Path Forward*, 6-3 (2009).
[18] Erica Beecher-Monas, *Evaluating Scientific Evidence: An Interdisciplinary Framework for Intellectual Due Process*, 97 (Cambridge Press 2007).
[19] Kennedy, *Forensic Science: Oxymoron*?, 302 Science 1625, 1625 (2003).
[20] NAS, *Strengthening Forensic Science in the United States: A Path Forward*, xx (2009).

needed in a number of forensic science disciplines—to ensure the reliability of the disciplines, establish enforceable standards, and promote best practices and their consistent application."[21]

10. "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."[22]

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."[23]

2. "Defense attorneys, prosecutors, judges and lay juries often lack scientific training and naively accept measurement results as certain."[24]

3. "[L]awyers…do not know how to think about validation of science claims."[25]

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."[26]

5. "[E]stablished case law in many jurisdictions supports minimal analytical quality control and documentation."[27]

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."[28]

7. "MANY LAWYERS SIMPLY COULD NOT DISTINGUISH BETWEEN REAL SCIENCE AND PRETENSIONS TO SCIENCE."[29]

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."[30]

---

[21] NAS, *Strengthening Forensic Science in the United States: A Path Forward*, xix (2009).
[22] NAS, *Strengthening Forensic Science in the United States: A Path Forward*, 3-2 (2009).
[23] NAS, *Strengthening Forensic Science in the United States: A Path Forward*, 3-20 (2009).
[24] Gullberg, *Estimating the measurement uncertainty in forensic breath-alcohol analysis*, 11 ACCRED. QUAL. ASSUR. 562, 563 (2006).
[25] 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).
[26] Bich, *Interdependence between measurement uncertainty and metrological traceability* ACCRED. QUAL. ASSUR. (IN PRESS - 2009).
[27] Gullberg, *Estimating the measurement uncertainty in forensic breath-alcohol analysis*, 11 ACCRED. QUAL. ASSUR. 562, 563 (2006).
[28] King, *Chemical measurement and the law: metrology and quality issues*, 6 ACCRED. QUAL. ASSUR. 236, 243 (2001).
[29] 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).

a. "Measurement…is the essential tool by which humans describe the world and reason about it."[31]

b. If "measurements are flawed, analyses and interpretations based on these measurements are fundamentally and irreparably fallacious."[32]

2. "MEASUREMENT THEORY IS THE CONCEPTUAL FOUNDATION OF ALL SCIENTIFIC DECISIONS."[33]

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."[34]

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."[35]

2. "…IF SCIENCE IS MEASUREMENT, THEN WITHOUT METROLOGY THERE CAN BE NO SCIENCE."[36]

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."[37]

---

30 Sir Isaac Newton, *Philosophiae Naturalis Principia Mathematica* (1687).
31 Finkelstein, *Expanding Technology, Deepening Knowledge and a Shrinking World: Reflections on Learned Societies in Measurement and Instrumentation*, 41 MEAS. CONTROL 170, 170 (2008).
32 Krebs, *Measurement Theory*, 67(12) PHYS. THERAPY 1834, 1839 (1987).
33 Krebs, *Measurement Theory*, 67(12) PHYS. THERAPY 1834, 1834 (1987).
34 JCGM, *International Vocabulary of Metrology — Basic and General Concepts and Associated Terms (VIM)*, § 2.2 (2008).
35 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).
36 William Thompson (Lord Kelvin), *Lecture to the Institution of Civil Engineers*, May 3, 1883.
37 U.S. Dept. of Labor, *Dictionary of Occupational Titles* 012.067-010.

8105 NE 140th Pl., Bothell WA 98011

B. BRIEF HISTORY:

1. Bible: "Have true scales, true weights and measures for all things."[38]

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."[39]

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 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."[40]

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.[41]

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.[42] Observation (examination) produces ***qualitative results*** indicating nominal and ordinal properties such as classification, identification and ordering.[43]

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.[44] I refer to them both under the common heading metrology herein for ease except where making a clear distinction is necessary.

---

38 Leviticus 19:36.
39 Magna Carta § 32.
40 John Quincy Adams, *Extract from the Report on Weights and Measures by the Secretary of State*, made to the Senate on February 22, 1821.
41 JCGM, *International Vocabulary of Metrology — Basic and General Concepts and Associated Terms (VIM)*, § 2.1 (2008).
42 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).
43 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).
44 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).

B. Weights & Measures

1. <u>Measurement Unit</u>: 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.[45]

   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 maintaining them as measurement standards, lies at the heart of fundamental metrology concepts and principles."[46]

   b. <u>The International System of Units (SI)</u>: 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.[47]

2. <u>Measurement Standard</u>: Realization of the definition of a given quantity, with stated quantity value and associated measurement uncertainty, used as a reference.[48]

   a. <u>Reference Material</u>: 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.[49]

      i. "The use of reference materials makes possible the transfer of the values of measured or assigned quantities between testing, analytical and measurement laboratories."[50]

      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."[51]

---

[45] JCGM, *International Vocabulary of Metrology — Basic and General Concepts and Associated Terms (VIM)*, § 1.9 (2008).
[46] *The Metrology Handbook* 149 (Bucher Ed. – 2004).
[47] JCGM, *The International System of Units (SI)* §1.2 (8th ed. 2008).
[48] JCGM, *International Vocabulary of Metrology — Basic and General Concepts and Associated Terms (VIM)*, § 5.1 (2008).
[49] 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).
[50] ISO, *General Requirements for the Competence of Reference Material Producers*, ISO Guide 34 v (2000).
[51] 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).

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."[52]

b. <u>Reference Procedure</u>: 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.[53]

C. <u>Measurement and Testing Process</u>:

1. <u>Subject</u>:

a. <u>Measurand</u>: Quantity intended to be measured.[54]

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."[55]

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."[56]

b. <u>Observand/Proto-Measurand</u>: "Particular nominal or ordinal property intended to be observed."[57]

i. <u>Ex.</u> "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."[58]

2. <u>Process</u>:

a. <u>Test Method</u>: Defined technical procedure to determine one or more specified characteristics of a material or product.[59]

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."[60]

---

[52] NIST, *Standard Reference Materials: Handbook for SRM Users*, NIST SP260-100, 16 (1993).
[53] JCGM, *International Vocabulary of Metrology — Basic and General Concepts and Associated Terms (VIM)*, § 2.7 (2008).
[54] JCGM, *International Vocabulary of Metrology — Basic and General Concepts and Associated Terms (VIM)*, § 2.3 (2008).
[55] JCGM, *International Vocabulary of Metrology — Basic and General Concepts and Associated Terms (VIM)*, § 2.3 Note 1 (2008).
[56] JCGM, *International Vocabulary of Metrology — Basic and General Concepts and Associated Terms (VIM)*, § 2.3 Note 2 (2008).
[57] Fuentes-Arderiu, *Vocabulary of terms in protometrology*, 11 Accred. Qual. Assur. 640, 642 (2006).
[58] JCGM, *International Vocabulary of Metrology — Basic and General Concepts and Associated Terms (VIM)*, § 2.3 Note 4 (2008).
[59] NIST, Handbook 150 § 1.5.29 (2006).

8105 NE 140th Pl., Bothell WA 98011

b. <u>Measurement Procedure (SOP)</u>: "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."[61]

c. <u>Measuring System</u>: 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.[62]

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)."[63]

3. <u>Validity</u>: "*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."[64]

a. <u>Validation</u>: "Validation is the confirmation by examination and the provision of objective evidence that the particular requirements for a specific intended use are fulfilled."[65]

i. "One particular task of science is the validation of new methods to determine their reliability under different conditions and their limitations."[66]

ii. ISO 17025 "includes a well established list of techniques that can be used, alone or in combination, to validate a method."[67]

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."[68]

---

[60] *The Metrology Handbook* 157 (Bucher Ed. – 2004).
[61] JCGM, *International Vocabulary of Metrology — Basic and General Concepts and Associated Terms (VIM)*, § 2.6 (2008).
[62] JCGM, *International Vocabulary of Metrology — Basic and General Concepts and Associated Terms (VIM)*, § 3.2 (2008).
[63] *The Metrology Handbook* 159-161 (Bucher Ed. – 2004).
[64] Krebs, *Measurement Theory*, 67(12) Phys. Therapy 1834, 1838 (1987).
[65] 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).
[66] NAS, *Strengthening Forensic Science in the United States: A Path Forward*, 113 (2009).
[67] NAS, *Strengthening Forensic Science in the United States: A Path Forward*, 113-114 (2009).
[68] 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."[69]

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."[70]

1) Validation of quantitative test methods must include statements of the uncertainty of the method such as documentation of precision and bias.[71]

2) "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."[72]

i. "Validation…establishes the crucial link between the metrological approach (analytical properties) and solving analytical problems (fitness for purpose)."[73]

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."[74]

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."[75]

---

[69] 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).
[70] 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).
[71] 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).
[72] Ellison, *Characterizing the performance of qualitative analytical methods: Statistics and terminology*, 24(6) TRENDS ANAL. CHEM. 468, 470 (2005).
[73] Rios, *Quality assurance of qualitative analysis in the framework of the European project 'MEQUALAN'*, 8 ACCRED. QUAL. ASSUR. 68, 74 (2003).
[74] NAS, *Strengthening Forensic Science in the United States: A Path Forward*, 114 (2009).
[75] 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).

1) "Commercial off-the-shelf software (e.g. word processing, database and statistical programs) in general use within their designed application range may be considered to be sufficiently validated. However, laboratory software configuration/modifications should be validated."[76]

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.'"[77]

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."[78]

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."[79]

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.[80]

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."[81]

b. Traceability includes the following essential elements:[82]

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;"

---

76 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).
77 Gullberg, *Methodology and Quality Assurance in Forensic Breath Alcohol Analysis*, 12 For. Sci. Rev. 49, 49 (2000).
78 EURACHEM, *The Fitness for Purpose of Analytical Methods A Laboratory Guide to Method Validation and Related Topics* § 4.4 (1998).
79 Shegunova, *Estimation of measurement uncertainty in organic analysis two practical approaches*, ACCRED. QUAL. ASSUR. (2008)
80 JCGM, *International Vocabulary of Metrology — Basic and General Concepts and Associated Terms (VIM)*, § 2.41 (2008); NIST, Handbook 150 § 1.5.30 (2006).
81 Hibbert, *Metrological traceability: I make it 42; you make it 42; but is it the same 42?* 13 ACCRED. QUAL. ASSUR. 11 (2006).
82 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."[83]

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."[84]

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."[85]

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."[86]

iii. "Comparability is an essential property of analytical results."[87]

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."[88]

[83] *The Metrology Handbook* 65 (Bucher Ed. – 2004).
[84] Hibbert, *Metrological traceability: I make it 42; you make it 42; but is it the same 42?* 11 ACCRED. QUAL. ASSUR. 543, 545 (2006).
[85] *Uncertainty Of Measurement In Biological, Forensic, Medical And Veterinary Testing*, NATA TECH. CIRC. 1 (December 2003).
[86] 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).
[87] Ellison, *Using validation data for ISO measurement uncertainty estimation Part 1. Principles of an approach using cause and effect analysis*, 123 ANALYST 1387 (1998).
[88] NIST, *Good Measurement Practice for Ensuring Traceability*, GMP-13, § 1.1 (2003).

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."[89]

iii. "The measurement of known and traceable standards is the basis for determining accuracy and thereby confidence in all analytical results."[90]

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."[91]

2. <u>CALIBRATION</u>: 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.[92]

   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."[93]

      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."[94]

      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."[95]

   b. <u>Uncertainty in Calibration</u>:

      i. Despite its importance, all "calibration…involves uncertainty."[96]

      ii. "A calibration is not complete until the expanded uncertainty associated with the calibration is determined and reported."[97]

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

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."[98]

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."[99]

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."[100]

ii. "It is not good measurement practice to report extrapolated data, i.e., outside the range calibrated."[101]

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."[102]

d. "Calibration with proper standards is the key to metrological traceability."[103]

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."[104]

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."[105]

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 in such a way that trends are detectable and, where practicable, statistical techniques shall be applied to the reviewing of the results."[106]

---

[97] NIST, *Good Laboratory Practice for Rounding Expanded Uncertainties and Calibration Values*, GLP-9, 1 (2003).
[98] NIST, *Standard Reference Materials: Handbook for SRM Users*, NISTSP 260-100, 6 (1993).
[99] NIST, *Good Measurement Practice for Standard Operating Procedure Selection*, GMP-12, 1 (2003).
[100] NIST, *Standard Reference Materials: Handbook for SRM Users*, NISTSP 260-100, 6 (1993).
[101] NIST, *Standard Reference Materials: Handbook for SRM Users*, NISTSP 260-100, 7 (1993).
[102] NIST, *Standard Reference Materials: Handbook for SRM Users*, NISTSP 260-100, 7 (1993).
[103] Hibbert, *Metrological traceability: I make it 42; you make it 42; but is it the same 42?* 11 ACCRED. QUAL. ASSUR. 543, 543 (2006).
[104] Rios, *Reliability of binary analytical responses*, 24(6) TRENDS ANAL. CHEM. 509, 510 (2005).
[105] Ellison, *Uncertainties in qualitative testing and analysis*, 5 ACCRED. QUAL. ASSUR. 346, 348 (2000).
[106] ISO, *General requirements for the competence of testing and calibration laboratories*, ISO 17025 § 5.9.1 (2005); NIST, Handbook 150 § 5.9.1 (2006).

a. <u>ACCREDITATION</u>: 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.[107]

i. "Accrediting bodies require that the methods meet a level of acceptable practice."[108]

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."[109]

ii. <u>Best Measurement Capability</u>: "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."[110]

iii. <u>Scope of Accreditation</u>: "The Scope of Accreditation lists the test methods or services, or calibration services, for which the laboratory is accredited."[111]

b. <u>PROFICIENCY TESTING</u>: Determination of laboratory testing performance by means of interlaboratory comparisons.[112]

i. "Proficiency testing requirements are associated with most fields of accreditation."[113]

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."[114] Types of processes subject to proficiency testing include:[115]

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

[107] NIST, Handbook 150 § 1.5.1 (2006); NAS, *Strengthening Forensic Science in the United States: A Path Forward*, 7-2 (2009).
[108] NAS, *Strengthening Forensic Science in the United States: A Path Forward*, 7-10 (2009).
[109] NIST, Handbook 150 App. B.2 (2006).
[110] NIST, Handbook 150 § 1.5.5 (2006).
[111] NIST, Handbook 150 § 1.5.26 (2006).
[112] NIST, Handbook 150 § 1.5.21 (2006).
[113] NIST, Handbook 150 § 3.4.2.1 (2006).
[114] NIST, Handbook 150 § 3.4.1.1 (2006).
[115] NIST, Handbook 150 § 1.5.21 (2006).

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."[116]

E. <u>MEASUREMENT INTERPRETATION</u>:

1. "It is scientific only to say what is more likely and what is less likely."[117]

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."[118]

2. <u>Measurement Result</u>: Set of quantity values being attributed to a measurand together with any other available relevant information.[119]

a. The value of a measurand can never be known exactly; all that can be known is its estimated value.[120]

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."[121]

a) <u>Ex</u>. 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:[122]

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

where
$\bar{y}$ = mean of measurements
ε = unknown uncertainty associated with mean

b. <u>UNCERTAINTY</u>: "Characterization of the dispersion of values assignable to a measurand based on the information available including systematic and random effects, definitional

[116] Gullberg, *Results of a Proposed Breath Alcohol Proficiency Test Program*, 51(1) J. For. Sci. 168,168 (2006).
[117] Feynman, *The Character of Physical Law* 165-166 (MIT Press 1965).
[118] NCCLS, Procedures for the Collection of Diagnostic Blood Specimens by Venipuncture; Approved Standard—Fifth Edition, H3-A5, § 5 (2003).
[119] JCGM, *International Vocabulary of Metrology — Basic and General Concepts and Associated Terms (VIM)*, § 2.9 (2008).
[120] 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).
[121] EURACHEM, *Guide to Quality in Analytical Chemistry* § 16.1 (2002).
[122] Eleftheriou, *Measuring performance in analytical measurements* 14 ACCRED. QUAL. ASSUR. 67, 67 (2009).

uncertainty and any other factors that may impact the measurement or test process or result."[123]

i. Uncertainty is a property of quantitative measurement results.[124]

ii. The estimate of uncertainty of a measurement:

a) "Quantifies the quality of a measurement result."[125]

b) "Reflects the lack of exact knowledge of the value of the measurand."[126]

c) "Is a necessary step in producing traceable results."[127]

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."[128]

3. OBSERVATION RESULT: "[E]stimated value of a particular nominal or ordinal property, obtained by observation."[129]

a. "Qualitative analysis is characterized by its binary nature: presence/absence, positive sample/negative sample, or yes/no according to a pre-set threshold." [130]

i. Types of qualitative analysis:[131]

a) Identification.

b) Classification.

---

[123] 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); Ted W. Vosk, *Uncertainty in Forensic Breath Alcohol Testing*, Intoxication Test Evidence, Ch. 56, (2nd Ed. 2009).
[124] Rios, *Quality assurance of qualitative analysis in the framework of the European project 'MEQUALAN'*, 8 ACCRED. QUAL. ASSUR. 68, 71 (2003).
[125] Croarkin, *Statistics and Measurements* 106 J. RES. NATL. INST. STAND. TECHNOL. 279, 283 (2001).
[126] JCGM, *Evaluation of measurement data — Guide to the expression of uncertainty in measurement (GUM)*, § 3.3.1 (2008).
[127] *Uncertainty Of Measurement In Biological, Forensic, Medical And Veterinary Testing*, NATA TECH. CIRC. 1 (December 2003).
[128] ISO, *Guidance for the use of repeatability, reproducibility and trueness estimates in measurement uncertainty estimation*, ISO/TS 21748 DRAFT REVISION, v (2009); Ted W. Vosk, *Uncertainty in Forensic Breath Alcohol Testing*, Intoxication Test Evidence, Ch. 56, (2nd Ed. 2009).
[129] Fuentes-Arderiu, *Vocabulary of terms in protometrology*, 11 ACCRED. QUAL. ASSUR. 640, 642 (2006).
[130] 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).
[131] Rios, *Quality assurance of qualitative analysis in the framework of the European project 'MEQUALAN'*, 8 ACCRED. QUAL. ASSUR. 68, 69 (2003).

8105 NE 140th Pl., Bothell WA 98011

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."[132] This is true "even when the analyst making the identification follows all the canons of best practice."[133]

iii. "Interpretation of the results must accordingly take the relevant uncertainties into account."[134]

iv. Uncertainty in qualitative methods is generally associated with the probabilistic determination of the reliability/unreliability of a method.[135]

b. <u>UNRELIABILITY</u>: The unreliability of a qualitative method is a measure of its likelihood of giving an erroneous response (error rate).[136]

i. "Traceability and (un)reliability of the [results] produced by these methods are crucial parameters in assuring the quality expected of the information derived."[137]

4. STATISTICAL APPROACHES:

a. <u>Frequentist (Relative frequency) Inference</u>: "To a frequentist the probability of an event is equal to its relative occurrence in a larger number of repetitions of an experiment."[138] In this approach, the population of past events sampled to determine a frequency of occurrence is assumed to be representative of the population of future events so that the frequency/probability found can be applied to the population future events.[139]

i. "[T]he only evidence employed in the decision making process are data which are derived from a sample."[140]

a) Probabilities are objectively determined as a function of empirical data.[141]

---

[132] 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).
[133] Ellison, *Quantifying uncertainty in qualitative analysis* 123 ANALYST 1155, 1155 (1998).
[134] Ellison, *Quantifying uncertainty in qualitative analysis* 123 ANALYST 1155, 1155 (1998).
[135] 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).
[136] 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).
[137] Rios, *Reliability of binary analytical responses*, 24(6) TRENDS ANAL. CHEM. 509, 515 (2005).
[138] Meinrath, *Lectures for chemists on statistics. I. Belief, probability, frequency, and statistics: decision making in a floating world*, 13 ACCRED. QUAL. ASSUR. 3, 7 (2008); Brüchle, *Confidence intervals for experiments with background and small numbers of events* 91 RADIOCHIM. ACTA 71, 71 (2003).
[139] Mendenhall, *Mathematical Statistics with Applications*, 17-18 (PWS-Kent 1990); *Handbook of Parametric and Nonparametric Statistical Procedures* 353 (CRC 2007).
[140] *Handbook of Parametric and Nonparametric Statistical Procedures* 332 (CRC 2007).

b) Uncertainty is treated using the concept of confidence intervals.[142]

ii. Drawbacks:[143]

a) Changed conditions: Relative frequencies or conditions for past events may not be the same as for future events.

b) Unique events: Cannot be applied to unique events.

b. Bayesian Inference: Probability is not a relative frequency of the occurrence of events, but an "information-based degree of belief about the truth of a proposition."[144] This approach combines subjective degrees of belief associated with relevant parameters prior to measurement/observation with the results of measurement/observation to determine updated degrees of belief incorporating new results.[145]

i. Bayesian inference employs sampling data and any other preexisting information deemed relevant in the decision making process.

a) Degrees of belief may be based upon both objective and subjective components.

ii. The foundation for Bayesian analysis is Bayes Theorem. It 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. This can be written as:[146]

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

where

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

---

141 *Handbook of Parametric and Nonparametric Statistical Procedures* 353 (CRC 2007).
142 JCGM, *International Vocabulary of Metrology — Basic and General Concepts and Associated Terms (VIM)*, § 2.36 Note 2 (2008); ISO, *Statistics — Vocabulary and symbols — Part 1: General statistical terms and terms used in probability*, ISO 3534-1, § 1.28 (2007).
143 *Handbook of Parametric and Nonparametric Statistical Procedures* 353 (CRC 2007).
144 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).
145 Bolstad, *Introduction to Baysian Statistics* 6-7 (Wiley 2007); Pearl, *Causality: Models Reasoning and Inference* 5-6 (Cambridge 2001); *Handbook of Parametric and Nonparametric Statistical Procedures* 355 (CRC 2007).
146 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).

a) "Bayesian inference provides a rigorous means of incorporating prior information into a measurement."[147]

iii. Drawbacks:

a) Validity of result is critically dependent upon the reliability of prior information and the validity of *subjectively* determined probabilities.[148]

c. The philosophy underlying each of these approaches is profoundly distinct.[149] The frequentist interpretation is that most widely espoused although Bayesian theory has gained prominence.[150] Regardless, both are important. 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, 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."[151]

b. MEASUREMENTS V. OBSERVATIONS:

i. "Traditional metrological principles, as they are applied to quantitative methods, cannot be directly applied to qualitative ones."[152] Accordingly quantitative and qualitative methods are treated separately.

---

[147] 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).
[148] Howson, *Scientific Reasoning The Bayesian Approach* 9 (Open Court 2006); Phillips, *Calculation of Measurement Uncertainty Using Prior Information* 103 J. RES. NATL. INST. STAND. TECHNOL. 625, 629 (1998); Ellison, *Quantifying uncertainty in qualitative analysis* 123 ANALYST 1155, 1160 (1998).
[149] 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).
[150] Croarkin, *Statistics and Measurements* 106 J. RES. NATL. INST. STAND. TECHNOL. 279, 290-291 (2001).
[151] ISO, *General requirements for the competence of testing and calibration laboratories*, ISO 17025 § 5.4.6.2 (2005).
[152] 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).

c. <u>UNCERTAINTY OF MEASUREMENT RESULTS (QUANTITATIVE METHODS)</u>:

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.*"[153]

ii. Basic Concepts:

a) <u>ACCURACY</u>: Closeness of agreement between a measured quantity value and a true quantity value of a measurand.[154]

1) Accuracy is not a quantity and is not given a numerical quantity value. A measurement is said to be more accurate when it offers a smaller measurement error.[155]

2) When multiple measures are made, accuracy is evaluated utilizing the mean of the set of measurements.

a) <u>ARITHMETIC MEAN</u>:[156] A sum of measurement values divided by the number of measurements.

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

b) <u>WEIGHTED MEAN</u>:[157] A 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.

$$\bar{y}_w = \frac{\sum_{i=1}^{n} w_i y_i}{\sum_{i}^{n} w_i}$$

where
ω_i = weighting factor

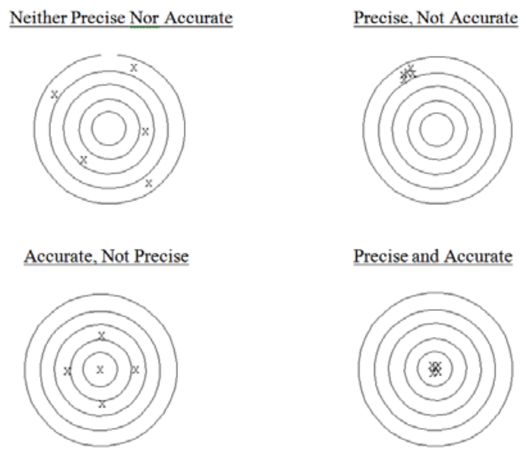
b) <u>PRECISION</u>: Closeness of agreement between indications or measured quantity values obtained by replicate measurements on the same or similar objects under specified conditions.[158]

[153] 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); Ted W. Vosk, *Uncertainty in Forensic Breath Alcohol Testing*, Intoxication Test Evidence, Ch. 56, (2nd Ed. 2009).
[154] JCGM, *International Vocabulary of Metrology — Basic and General Concepts and Associated Terms (VIM)*, § 2.13 (2008).
[155] JCGM, *International Vocabulary of Metrology — Basic and General Concepts and Associated Terms (VIM)*, § 2.13 Note 1 (2008).
[156] ISO, *Statistics — Vocabulary and symbols — Part 1: General statistical terms and terms used in probability*, ISO 3534-1, § 1.15 (2007).
[157] 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).
[158] JCGM, *International Vocabulary of Metrology — Basic and General Concepts and Associated Terms (VIM)*, § 2.15 (2008).

1) MEASUREMENT STANDARD DEVIATION:[159] Characterizes the variability/dispersion of measured values about their mean.

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

c) A set of measurements may be neither accurate nor precise, precise but not accurate, accurate but not precise or both accurate and precise.[160]



1) "Accuracy…is judged with respect to the use to be made of the data."[161]

2) "What might be considered as very precise for one purpose could be grossly imprecise for another."[162]

d) MEASUREMENT ERROR: Measured quantity value minus a reference quantity value.[163] "Traditionally, an error is viewed as having two components, namely, a random component and a systematic component."[164]

1) SYSTEMATIC ERROR: Component of measurement error that in replicate measurements remains constant or varies in a predictable manner.[165]

a) BIAS: Estimate of a systematic error.[166]

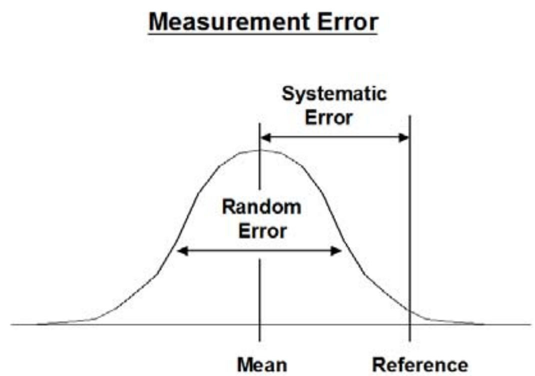
---

159 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).
160 Dimech, *Calculating Uncertainty of Measurement for Serology Assays by Use of Precision and Bias* 52(3) CLIN. CHEM. 526, 527 (2006).
161 NIST, *Standard Reference Materials: Handbook for SRM Users*, NISTSP 260-100, 2 (1993).
162 NIST, *Standard Reference Materials: Handbook for SRM Users*, NISTSP 260-100, 3 (1993).
163 JCGM, *International Vocabulary of Metrology — Basic and General Concepts and Associated Terms (VIM)*, § 2.16 (2008).
164 JCGM, *Evaluation of measurement data — Guide to the expression of uncertainty in measurement (GUM)*, § 3.2.1 (2008).
165 JCGM, *International Vocabulary of Metrology — Basic and General Concepts and Associated Terms (VIM)*, § 2.17 (2008).
166 JCGM, *International Vocabulary of Metrology — Basic and General Concepts and Associated Terms (VIM)*, § 2.18 (2008).

1) "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."[167]

2) <u>RANDOM ERROR</u>: Component of measurement error that in replicate measurements varies in an unpredictable manner.[168]

[169]



3) "Error analysis is the attempt to estimate the total error using frequency-based statistics."[170]

iii. <u>CHARACTERIZING ACCURACY</u>: "Accuracy…includes the concepts of both bias and precision and is <u>*judged with respect to the use to be made of the data*</u>. 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."[171]

---

[167] NIST, *Standard Reference Materials: Handbook for SRM Users*, NISTSP 260-100, 4 (1993).
[168] JCGM, *International Vocabulary of Metrology — Basic and General Concepts and Associated Terms (VIM)*, § 2.19 (2008).
[169] Image by Rod Gullberg.
[170] Ehrlich, *Evolution of philosophy and description of measurement* 12 ACCRED. QUAL. ASSUR. 201, 205 (2007).
[171] NIST, *Standard Reference Materials: Handbook for SRM Users*, NISTSP 260-100, 2 (1993).

iv. <u>BEST ESTIMATE OF MEASURAND VALUE</u>:

a) "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."[172]

b) "It is understood that the result of the measurement is the best estimate of the value of the measurand, and that all components of uncertainty, including those arising from systematic effects, such as components associated with corrections and reference standards, contribute to the dispersion."[173]

c) When multiple measurements are obtained, the best estimate of Y may be based on either an arithmetic or weighted mean.[174] Although in special circumstances the weighted and classical mean may be equal, in general they will not be.[175]

1) <u>Within Laboratory Measurements</u>: 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."[176]

a) Arithmetic mean is appropriate.

2) <u>Between Laboratory Measurement</u>: Some measurements performed where either method, conditions, analysts, operators, instruments or laboratories are different.[177]

a) Weighted mean accepted approach.[178]

---

[172] JCGM, *International Vocabulary of Metrology — Basic and General Concepts and Associated Terms (VIM)*, § 0.1 (2008); Ted W. Vosk, *Uncertainty in Forensic Breath Alcohol Testing*, Intoxication Test Evidence, Ch. 56, (2nd Ed. 2009).
[173] JCGM, *Evaluation of measurement data — Guide to the expression of uncertainty in measurement (GUM)*, § 2.2.3 Note 3 (2008).
[174] 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).
[175] Paule, *Consensus Values and Weighting Factors*, 87 J. RES. NAT'L BUREAU STAND. 377, 380 (1982).
[176] Jones, *Dealing with Uncertainty in Chemical Measurements*, 14(1) NEWSL. OF THE INT. ASSOC. FOR CHEM. TEST. 8 (2003).
[177] 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).
[178] 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*

1) 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: [179]

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

2) 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.[180]

3) Under the principle of maximum likelihood, the weighted mean yields the most precise value for the best estimate of Y.[181]

4) Failure to utilize the weighted mean in these circumstances can result in an underestimation of uncertainty.[182]

5) "There are many situations in which it would be very misleading to average quantities without [weighting them]".[183]

v. 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 appropriate methods of analysis."[184]

b) All known systematic effects should be compensated for through application of a correction factor.[185] "It is assumed that a correction (or correction factor) is applied to compensate for each recognized systematic effect that significantly

Users, NIST SP260-100, 78 (1993); ISO, *Reference Materials – General and Statistical Principles for Certification*, ISO Guide 35, App. B.7 (2006).

[179] 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).

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

[181] 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).

[182] 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).

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

[184] ISO, *General requirements for the competence of testing and calibration laboratories*, ISO 17025 § 5.4.6.3 (2005).

[185] 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).

influences the measurement result."[186] Assuming we have determined a systematic error (bias) of b, our best estimate of Y would then be:[187]

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

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."[188]

1) <u>Type A Uncertainty</u>: Component of measurement uncertainty determined by a statistical analysis of a series of measured quantity values obtained under defined measurement conditions.[189]

a) "A Type A evaluation of standard uncertainty may be based on any valid statistical method for treating data."[190]

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

2) 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;

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

2) <u>Type B Uncertainty</u>: Component of measurement uncertainty determined by a method other than the statistical analysis of series of observations.[191]

a) "A Type B evaluation of standard uncertainty is usually based on scientific judgment using all the relevant information available."[192]

1) Previous measurement data;

---

[186] NIST, *Guidelines for Evaluating and Expressing the Uncertainty of NIST Measurement Results*, NIST TN 1297, §5.2, App. D 1.1.6 – 8 (1994).
[187] Eleftheriou, *Measuring performance in analytical measurements* 14 Accred. Qual. Assur. 67, 67 (2009); Ted W. Vosk, *Uncertainty in Forensic Breath Alcohol Testing*, Intoxication Test Evidence, Ch. 56, (2nd Ed. 2009).
[188] NIST, Handbook 150 § 1.5.31 Note 2 (2006).
[189] 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).
[190] NIST, *Guidelines for Evaluating and Expressing the Uncertainty of NIST Measurement Results*, NIST TN 1297, § 3 (1994).
[191] 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).
[192] 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).

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

3) Manufacturer's specifications;

4) Data provided in calibration and other reports;

5) Uncertainties assigned to reference data taken from handbooks;

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

7) Information about instrumental drift.

3) "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."[193]

a) "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."[194]

b) "[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."[195]

1) "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."[196]

2) "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."[197]

d) "Sources contributing to the uncertainty include, but are not necessarily limited to, the reference standards and reference materials used, methods and equipment

[193] JCGM, *Evaluation of measurement data — Guide to the expression of uncertainty in measurement (GUM)*, § 3.3.4 (2008).
[194] JCGM, *Evaluation of measurement data — Guide to the expression of uncertainty in measurement (GUM)*, § 4.1.6 (2008).
[195] Ehrlich, *Evolution of philosophy and description of measurement* 12 ACCRED. QUAL. ASSUR. 201, 212-213 (2007).
[196] 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).
[197] Ehrlich, *Evolution of philosophy and description of measurement* 12 ACCRED. QUAL. ASSUR. 201, 213 (2007).

8105 NE 140th Pl., Bothell WA 98011

used, environmental conditions, properties and condition of the item being tested or calibrated, and the operator."[198]

e) <u>UNCERTAINTY BUDGET</u>: Statement of a measurement uncertainty, of the components of that measurement uncertainty, and of their calculation and combination.[199]

[200]

Source of uncertainty	Type A	Type B
1. Factors influencing only the reference standard:		
1.1 Constancy of the dosimeter		0.100
1.2 Dosimeter reading	0.010	0.010
1.3 Temperature *T*: difference with *T* inside the cavity		0.060
Thermometer resolution		0.020
1.4 Pressure		0.060
1.5 Current/charge measurements		0.060
1.6 Reproducibility of the phantom positioning	0.040	
Quadratic sum	0.041	0.144
Combined uncertainty 1	0.150	
2. Factors influencing only the user's chamber		
2.1 Dosimeter reading	0.020	0.060
2.2 Temperature: difference with *T* inside the cavity		0.060
Thermometer resolution		0.020
2.3 Pressure		0.060
2.4 Current/charge measurements		0.060
2.5 Leakage current		0.020
2.6 Reproducibility of the phantom positioning	0.040	
Quadratic sum	0.045	0.108
Combined uncertainty 2	0.117	
3. Total uncertainty		
3.1 Quadratic sum (1 + 2)	0.061	0.180
3.2 Combined uncertainty SSDL (1 + 2)	0.190	
3.3 Uncertainty of the calibration coefficient reported by IAEA		0.490
3.4 Combined uncertainty (SSDL + IAEA)	0.526	
3.5 Expanded uncertainty ($k = 2$)	1.052	

f) <u>COMBINED UNCERTAINTY</u>: Measurement uncertainty determined by combining the uncertainties associated with each individual source of uncertainty identified

[198] 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).

[199] JCGM, *International Vocabulary of Metrology — Basic and General Concepts and Associated Terms (VIM)*, § 2.33 (2008); Ted W. Vosk, *Uncertainty in Forensic Breath Alcohol Testing*, Intoxication Test Evidence, Ch. 56, (2nd Ed. 2009).

[200] Image from Arib, *Study of the influence of phantom material and size on the calibration of ionization chambers in terms of absorbed dose to water* 7(3) J. APP. CLIN. MED. PHYS. (2006) at www.jacmp.org/index.php/jacmp/article/viewArticle/2264/1286.

in the uncertainty budget.[201] The combined uncertainty can be represented symbolically as: μ_c

g) <u>EXPANDED UNCERTAINTY</u>: Value obtained when the combined uncertainty is multiplied by a "coverage factor." The expanded uncertainty can be represented symbolically as: $U = \lambda\mu_c$[202]

1) The expanded uncertainty defines a "coverage interval" that may be expected to encompass the set of values that could reasonably be attributed to the measurand with a stated probability based on the information available.[203]

2) <u>LEVEL OF CONFIDENCE</u>: Probability that the set of true quantity values of a measurand is contained within a specified coverage interval.[204]

a) 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.[205]

b) For a given set of assumptions, the level of confidence provided by an interval is determined by the coverage factor chosen.[206]

h) THE COVERAGE INTERVAL:

1) A coverage interval need not be symmetric with respect to (centered on) the chosen measured quantity value.[207]

---

[201] 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).

[202] 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).

[203] 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).

[204] 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).

[205] JCGM, *Evaluation of measurement data — Guide to the expression of uncertainty in measurement (GUM)*, § 2.3.5 (2008).

[206] JCGM, *Evaluation of measurement data — Guide to the expression of uncertainty in measurement (GUM)*, § 6.3.1 (2008).

[207] 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).

2) Frequently (although by no means universally) measurement variables are approximately normally distributed. In these circumstances, the coverage interval will be symmetric about the corrected mean of our measurement results so that our best estimate of Y becomes:[208]

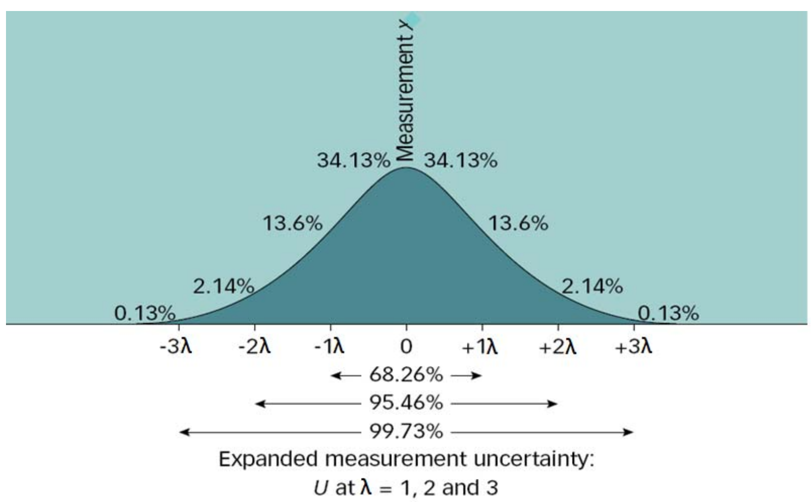
$$\begin{aligned} Y &= y \pm U \\ &= y \pm \lambda\mu_c \end{aligned}$$

a) This "is interpreted to mean that the best estimate of the value attributable to the measurand Y is y, and that $y - U$ to $y + U$ is an interval that may be expected to encompass a large fraction of the distribution of values that could reasonably be attributed to Y."[209]

b) In this context, our level of confidence is determined by our coverage factor, λ.

c) Setting $\lambda = 2$ produces an interval having a level of confidence of approximately 95 percent, while setting $\lambda = 3$ produces an interval having a level of confidence of approximately 99 percent.[210]

[211]



[208] JCGM, *Evaluation of measurement data — Guide to the expression of uncertainty in measurement (GUM)*, § 6.2.1 (2008); 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).

[209] JCGM, *Evaluation of measurement data — Guide to the expression of uncertainty in measurement (GUM)*, § 6.2.1 (2008).

[210] JCGM, *Evaluation of measurement data — Guide to the expression of uncertainty in measurement (GUM)*, § 3.3.7, 6.3.3 (2008); NIST, *Guidelines for Evaluating and Expressing the Uncertainty of NIST Measurement Results*, NIST TN 1297 § 6.2 (1994); UKAS, *The Expression of Uncertainty and Confidence in Measurement*, M3003 § 3.47 (2007).

[211] Image adapted from, Shah, *Standard Definition Getting to the bottom of measurement uncertainty* 42(3) QUALITY PROGRESS 53 (2009).

3) Although a confidence interval is an example of a coverage interval, a coverage interval is not a confidence interval. A confidence interval is a statistical concept based on the frequentist approach while a coverage interval is a metrological concept. Although conceptually similar, only when certain assumptions are satisfied will a coverage interval be a confidence interval.[212]

a) In the context of a confidence interval, "[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. A confidence interval does not reflect the probability that the observed interval contains the true value of the parameter (it either does or does not contain it)."[213]

i) SAFETY MARGIN:

1) 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."[214]

2) 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."[215]

3) This is similar to utilization of a one sided confidence interval.[216] Assuming a safety margin s, our estimate of Y becomes:

$$Y \geq y - s$$

4) 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 $Y < y - s$.[217]

---

[212] JCGM, *International Vocabulary of Metrology — Basic and General Concepts and Associated Terms (VIM)*, § 2.36 (2008); JCGM, *Evaluation of measurement data — Guide to the expression of uncertainty in measurement (GUM)*, § 6.2.2 (2008); ISO, *Statistics — Vocabulary and symbols — Part 1: General statistical terms and terms used in probability*, ISO 3534-1, § 1.28 (2007).
[213] ISO, *Statistics — Vocabulary and symbols — Part 1: General statistical terms and terms used in probability*, ISO 3534-1, § 1.28 (2007); Ted W. Vosk, *Uncertainty in Forensic Breath Alcohol Testing*, Intoxication Test Evidence, Ch. 56, (2nd Ed. 2009).
[214] Kristiansen, *An Uncertainty Budget for the Measurement of Ethanol in Blood by Headspace Gas Chromatography*, 28(6) J. ANAL. TOX. 456 (2004).
[215] Kristiansen, *An Uncertainty Budget for the Measurement of Ethanol in Blood by Headspace Gas Chromatography*, 28(6) J. ANAL. TOX. 456 (2004).
[216] ISO, *Statistics — Vocabulary and symbols — Part I: General statistical terms and terms used in probability*, ISO 3534-1 § 1.29 (2006).
[217] 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. <u>UNRELIABILITY OF OBSERVATIONS (QUALITATIVE METHODS)</u>:

i. FREQUENTIST METHODS: Traditional frequentist probabilistic measures of unreliability include:[218]

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."[219]

[218] 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).

[219] NAS, *Strengthening Forensic Science in the United States: A Path Forward*, 4-9 (2009).

h) When utilizing frequentist methods, "[i]t is important for laboratories to check at least the most critical false response rate for a qualitative test."[220]

ii. <u>BAYESIAN METHODS</u>: An important alternative to frequentist measures of unreliability is the application of Bayesian measures.[221]

a) <u>LIKELIHOOD RATIO</u>: [222] $L(I \mid H) = \frac{p(I \mid H)}{p(I \mid \neg H)}$. 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.

b) <u>POSTERIOR PROBABILITY</u>:[223] $p(H \mid I)$ = Probability (degree of belief) that H is true given test result I.

c) <u>BAYESIAN NETWORK</u>: "Bayesian networks are a graphical representation of (in)dependencies amongst random variables…with nodes representing random variables, and arcs representing direct influence… Bayesian networks aid in knowledge acquisition by specifying which probabilities are needed."[224]

6. <u>REPORTING RESULTS</u>:

a. "Calculations and data transfers shall be subject to appropriate checks in a systematic manner."[225]

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."[226]

ii. <u>OUTLIER</u>: "[A]n observation that appears to deviate markedly in value from other members of the sample in which it appears."[227]

---

[220] Ellison, *Uncertainties in qualitative testing and analysis*, 5 ACCRED. QUAL. ASSUR. 346, 348 (2000).

[221] 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).

[222] 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).

[223] 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).

[224] Zhang, *Exploiting Causal Independence in Bayesian Network Inference* 5 J. ART. INTEL. RES. 301, 301 (1996).

[225] 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).

[226] ASTM, *Standard Practice for Dealing With Outlying Observations*, E 178 § 4.1 (2008).

[227] 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).

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."[228]

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.[229]

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."[230]

1) "Outliers should not be excluded on purely statistical evidence until they have been thoroughly investigated and, where possible, the reasons for the discrepancies identified."[231]

2) Chauvenet's Criterion (also known as Grubb's test) is a common test for outliers:[232]

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

a) The value chosen for C determines the level of confidence of the outlier test.[233]

d) "For qualitative methods, statistical outliers are represented by abnormally high or low frequencies of incorrect responses."[234]

b. RESULT = MEASUREMENT + UNCERTAINTY:

i. "The result of a measurement cannot be correctly evaluated without knowing its uncertainty."[235]

---

[228] 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).
[229] 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).
[230] ISO, *Reference Materials – General and Statistical Principles for Certification*, ISO Guide 35 § 10.5.5 (2006).
[231] 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).
[232] 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).
[233] 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).
[234] Ellison, *Characterizing the performance of qualitative analytical methods: Statistics and terminology*, 24(6) TRENDS ANAL. CHEM. 468, 475 (2005).

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."[236]

ii. "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 be compared, either among themselves or with reference values given in a specification or standard."[237]

a) "Requirements for measurement accuracy translate into a need to know not only the results of measurements but the uncertainties associated with the results."[238]

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."[239]

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."[240]

iii. <u>Reports of result must include</u>:[241]

a) <u>Test method</u> – Description of how test was made;

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

c) <u>Standards used</u> – Identification of and traceability to national standards;

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

---

235 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); Ted W. Vosk, *Uncertainty in Forensic Breath Alcohol Testing*, Intoxication Test Evidence, Ch. 56, (2nd Ed. 2009).
236 Watson, *Pharmaceutical Analysis - A Textbook for Pharmacy Students and Pharmaceutical Chemists*, 2 (2nd ed. Elsevier 2005); Ted W. Vosk, *Uncertainty in Forensic Breath Alcohol Testing*, Intoxication Test Evidence, Ch. 56, (2nd Ed. 2009).
237 JCGM, *Evaluation of measurement data — Guide to the expression of uncertainty in measurement (GUM)*, § 0.1 (2008).
238 Ehrlich, *Metrological Timelines in Traceability*, 103 J. Res. Natl. Inst. Stand. Technol. 93, 94 (1998).
239 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).
240 Richter, *Reporting measurement uncertainty in chemical analysis*, 13 ACCRED. QUAL. ASSUR. 113, 113 (2008).
241 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).

1) Include all corrections and constants used in the analysis and their sources;

e) Estimated measurement uncertainty –

1) List all uncertainty components and document fully how they were evaluated;

2) Coverage factor and estimated confidence interval.

iv. 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."[242]

b) COVERAGE INTERVAL APPROACH:

1) "State the result of the measurement as $Y = y \pm U$ and give the units of y and U."[243]

2) Give the value of λ used to obtain U ($U = \lambda\mu_c$).[244]

3) "Give the approximate level of confidence associated with the interval $y \pm U$ and state how it was determined."[245]

4) The expanded uncertainty is generally reported with approximately a 95% - 99% level of confidence.[246]

c) SAFETY MARGIN APPROACH:

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

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

c. RESULTS OF QUALITATIVE TEST OBSERVATIONS:

i. FREQUENTIST APPROACH: "The most common, and probably the most useful, form of data treatment in method-validation studies for qualitative tests is the calculation and

---

[242] 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); Ted W. Vosk, *Uncertainty in Forensic Breath Alcohol Testing*, Intoxication Test Evidence, Ch. 56, (2nd Ed. 2009).
[243] JCGM, *Evaluation of measurement data — Guide to the expression of uncertainty in measurement (GUM)*, § 7.2.3 (2008).
[244] JCGM, *Evaluation of measurement data — Guide to the expression of uncertainty in measurement (GUM)*, § 7.2.3 (2008).
[245] JCGM, *Evaluation of measurement data — Guide to the expression of uncertainty in measurement (GUM)*, § 7.2.3 (2008).
[246] 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).

reporting of either specificity and sensitivity or false positive and negative error rates."[247]

ii. BAYESIAN APPROACH: "The scientist can testify to the value of their evidence by quoting a likelihood ratio value obtained from a particular procedure."[248]

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."[249]

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."[250]

a. "Standards should be based on the consolidated results of science, technology and experience, and aimed at the promotion of optimum community benefits."[251]

b. Types

i. Basic Standard: Standard that has a wide-ranging coverage or contains general provisions for one particular field.[252]

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.[253]

iii. Process Standard: Standard that specifies requirements to be fulfilled by a process, to establish its fitness for purpose.[254]

iv. Terminology Standard: Define words permitting parties to use a common, clearly understood language.[255]

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.[256]

---

[247] Ellison, *Characterizing the performance of qualitative analytical methods: Statistics and terminology*, 24(6) TRENDS ANAL. CHEM. 468, 470 (2005).
[248] 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).
[249] Eleftheriou, *Measuring performance in analytical measurements* 14 ACCRED. QUAL. ASSUR. 67, 67 (2009).
[250] ISO, *Standardization and related activities — General vocabulary*, ISO 2 § 3.2 (2004).
[251] ISO, *Standardization and related activities — General vocabulary*, ISO 2 § 3.2 Note (2004).
[252] ISO, *Standardization and related activities — General vocabulary*, ISO 2 § 5.1 (2004).
[253] ISO, *Standardization and related activities — General vocabulary*, ISO 2 § 5.3 (2004).
[254] ISO, *Standardization and related activities — General vocabulary*, ISO 2 § 5.5 (2004).
[255] 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).

c. ACKNOWLEDGED RULE OF TECHNOLOGY: Technical provision acknowledged by a majority of representative experts as reflecting the state of the art.[257]

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.[258]

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."[259]

iii. "Voluntary consensus standards are heavily peer-reviewed before they even come into existence."[260]

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."[261]

ii. "Standards ensure desirable characteristics of services and techniques such as quality, reliability, efficiency, and consistency among practitioners."[262]

iii. "[S]tandards are crucial to every form of scientific and industrial process."[263]

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."[264]

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

[256] ISO, *Standardization and related activities — General vocabulary*, ISO 2 § 5.8 (2004).
[257] ISO, *Standardization and related activities — General vocabulary*, ISO 2 § 1.5 (2004).
[258] ISO, *Standardization and related activities — General vocabulary*, ISO 2 § 1.4 (2004).
[259] ISO, *Standardization and related activities — General vocabulary*, ISO 2 § 1.5 Note (2004).
[260] Lentini, *Forensic Science Standards: Where They Come From and How They Are Used*, 1 FOR. SCI. POL. MGMT. 10, 10 (2009).
[261] NAS, *Strengthening Forensic Science in the United States: A Path Forward*, 7-7 (2009).
[262] NAS, *Strengthening Forensic Science in the United States: A Path Forward*, 7-1 (2009).
[263] Breitenberg, Office of Standards Code and Information, NIST, *The ABC's of Standards-Related Activities in the United States*, NBSIR 87-3576 (1987).
[264] NAS, *Strengthening Forensic Science in the United States: A Path Forward*, 7-1 – 7-2 (2009).

calibration performed using standard methods, non-standard methods, and laboratory-developed methods."[265]

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.[266]

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."[267]

c. "This international standard forms the basis for international laboratory accreditation."[268]

V. FORENSIC METROLOGY

A. FORENSIC METROLOGY: "Forensic Metrology is the application of measurements and hence measurement standards to the solution and prevention of crime."[269]

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."[270]

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."[271]

3. Forensic metrology is practiced around the world.[272]

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"[273] as well as to assure the validity of qualitative test results.[274]

---

265 ISO, *General requirements for the competence of testing and calibration laboratories*, ISO 17025 § 1.1 (2005).
266 NIST, Handbook 150 § 1.5.8 (2006).
267 ISO, *General requirements for the competence of testing and calibration laboratories*, ISO 17025 § 1.2 (2005); NIST, Handbook 150, *v-vi* (2006).
268 UKAS, *The Expression of Uncertainty and Confidence in Measurement*, M3003 § 1.1 (2007).
269 Sharp, *Measurement Standar*ds, in *Measurement, Instrumentation, and Sensors Handbook* §5.2 (1999).
270 King, *Chemical measurement and the law: metrology and quality issues*, 6 ACCRED. QUAL. ASSUR. 236, 241 (2001).
271 Kaarls, *Metrology, essential to trade, industry and society*, 12 ACCRED. QUAL. ASSUR. 423, 435 (2007).
272 Sharp, *Measurement Standar*ds, in *Measurement, Instrumentation, and Sensors Handbook* §5.2 (1999).
273 NIST, *1999 Survey of Forensic Reference Materials*, NISTIR 6518, 1 (2000).

b. "Appropriate reference material(s) shall be used for qualitative and quantitative procedures. Traceability of the reference material is required."[275]

i. <u>CERTIFIED REFERENCE MATERIAL</u>: 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.[276]

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."[277]

ii. <u>REFERENCE STANDARD</u>: 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.[278]

c. <u>Adequacy and documentation of references</u>:

i. "Clear documentation of the [reference material] and its property value(s) should be available, preferably as a certificate ([certified reference material])."[279]

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 date and or expiry date; identity of preparer; storage conditions, if relevant; hazard warning, where necessary."[280]

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."[281]

---

[274] 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).
[275] SWGDRUG, *Recommendations (Minimum Standards)*, 31 (2008).
[276] 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).
[277] SOFT/AAFS, *Forensic Toxicology Laboratory Guidelines*, § 9.3.1 (2006).
[278] ASCLD/LAB – International, *Traceability Discussion*, 2 (2004).
[279] Epstein, *The Use of Certified Reference Materials in Forensic QA*, Presented at 13th INTERPOL Forensic Science Symposium, (2001).
[280] ILAC, *Guideline for Forensic Science Laboratories*, ILAC G19, § 5.4.2d (2002).
[281] ILAC, *Guideline for Forensic Science Laboratories*, ILAC G19, § 5.6.3.2 (2002).

d. "[L]aboratories may obtain certified reference material from NIST…or from another national metrology institute."[282]

C. FORENSIC MEASUREMENT AND TESTING PROCESS

1. TEST METHOD: Defined technical procedure to determine one or more specified characteristics of a material or product.[283]

a. "All methods shall be fully documented including procedures for quality control, and, where appropriate, the use of reference materials."[284]

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."[285]

a. "Visual inspection, qualitative examinations and computer simulations are included in the definition of objective test."[286]

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."[287]

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."[288]

3. VALIDATION: Confirmation by examination and provision of objective evidence that the particular requirements for a specific intended use are fulfilled.[289]

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."[290]

b. "All technical procedures used by a forensic science laboratory must be fully validated before being used on casework."[291]

---

[282] 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).
[283] ASTM, *Standard Terminology Relating to Forensic Science*, §4 E 1732 (2005).
[284] FQS-I, *Forensic Requirements for Accreditation*, FRA-1, § 5.4.1 (2008).
[285] ILAC, *Guideline for Forensic Science Laboratories*, ILAC G19, § 3 (2002).
[286] ILAC, *Guideline for Forensic Science Laboratories*, ILAC G19, § 3 (2002).
[287] FQS-I, *Forensic Requirements for Accreditation*, FRA-1, § 3 (2008).
[288] ILAC, *Guideline for Forensic Science Laboratories*, ILAC G19, § 3 (2002).
[289] ASTM, *Standard Terminology Relating to Forensic Science*, §4 E 1732 (2005); Ted Vosk, *The DataMaster*, Defending DUIs in Washington Ch.13 (3rd Ed. 2008).
[290] NAS, *Strengthening Forensic Science in the United States: A Path Forward*, 113 (2009).

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."[292]

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."[293]

a. "The contribution of random and systematic errors to method result uncertainty shall be assessed and the expanded uncertainty derived for quantitative methods."[294]

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."[295]

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."[296]

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).[297]

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."[298]

ii. "Especially for legal purposes, traceability is an essential requirement, which however is not always fulfilled."[299]

---

291 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).
292 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).
293 Gullberg, *Estimating the measurement uncertainty in forensic breath-alcohol analysis*, 11 ACCRED. QUAL. ASSUR. 562, 562 (2006).
294 SWGDRUG, *Recommendations (Minimum Standards)*, 34 (2008).
295 ILAC, *Guideline for Forensic Science Laboratories*, ILAC G19, § 5.4.5.1 (2002).
296 Gullberg, *Methodology and Quality Assurance in Forensic Breath Alcohol Analysis*, 12 For. Sci. Rev. 49, 56 (2000).
297 ASTM, *Standard Terminology Relating to Forensic Science*, §4 E 1732 (2005).
298 Knopf, *Traceability system for breath-alcohol measurements in Germany*, OIML Bulletin XLVIII(2), 17 (2007).
299 Bich, *Interdependence between measurement uncertainty and metrological traceability* ACCRED. QUAL. ASSUR. (IN PRESS - 2009).

iii. "Traceability to authoritative reference standards is an important and often overlooked element in forensic…analysis."[300]

iv. "[B]ias can be corrected when traceability is established."[301]

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."[302]

i. "ISO/IEC 17025 details the specific requirements for traceablilty to be met by testing and calibration laboratories."[303]

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.[304]

c. <u>DOCUMENTATION</u>: "Accounting for and documenting traceability…is an important element of quality control."[305]

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."[306]

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 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."[307]

iii. "The uncertainty of measurement for each step in the traceability chain must be determined and stated."[308]

---

[300] Gullberg, *Methodology and Quality Assurance in Forensic Breath Alcohol Analysis*, 12 For. Sci. Rev. 49, 59 (2000).
[301] Gullberg, *Estimating the measurement uncertainty in forensic breath-alcohol analysis*, 11 ACCRED. QUAL. ASSUR. 562, 563 (2006).
[302] FQS-I, *Traceability*, FRAP-4 (2008).
[303] FQS-I, *Traceability*, FRAP-4 (2008).
[304] FQS-I, *Traceability*, FRAP-4 § 2.2 (2008).
[305] Gullberg, *Estimating the measurement uncertainty in forensic breath-alcohol analysis*, 11 ACCRED. QUAL. ASSUR. 562, 568 (2006).
[306] ASCLD/LAB – International, *Measurement Traceability Policy*, 1 (2004).
[307] ASCLD/LAB – International, *Measurement Traceability Policy*, 1 (2004).
[308] ASCLD/LAB – International, *Traceability Discussion*, 2 (2004).

2. <u>CALIBRATION</u>: 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.[309]

   a. <u>When required</u>:

      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."[310]

      ii. "Calibration must be performed…at appropriate intervals thereafter."[311]

      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."[312]

   b. <u>Calibration defines the valid range of measurement</u>:

      i. "The concentration of the calibrators should be such that they bracket the anticipated concentration of the specimen(s)."[313]

      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."[314]

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 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."[315]

   b. <u>ACCREDITATION</u>: "[P]rocedure by which an authoritative body gives formal recognition that a body or person is competent to carry out specific tasks."[316]

---

[309] ASTM, *Standard Terminology Relating to Forensic Science*, §4 E 1732 (2005).
[310] ASCLD/LAB – International, *Traceability Discussion*, 1 (2004).
[311] ASCLD/LAB – International, *Traceability Discussion*, 2 (2004).
[312] ILAC, *Guideline for Forensic Science Laboratories*, ILAC G19, § 5.6.1 (2002).
[313] SOFT/AAFS, *Forensic Toxicology Laboratory Guidelines*, § 8.3.6 (2006).
[314] Flanagan, *Fundamentals of Analytical Toxicology* 357 (Wiley 2007).
[315] NAS, *Strengthening Forensic Science in the United States: A Path Forward*, 7-19 (2009); Ted Vosk, *The DataMaster*, Defending DUIs in Washington Ch.13 (3rd Ed. 2008).
[316] ASTM, *Standard Terminology Relating to Forensic Science*, E 1732 § 4.1 (2005).

i. "Accreditation deals directly with the ability of a laboratory to provide quality forensic science service."[317]

ii. Accreditation of forensic laboratories must be mandatory.[318]

iii. "Accreditation is part of a laboratory's quality assurance program which should also include proficiency testing."[319]

c. PROFICIENCY TESTING:

i. Proficiency testing is an important aspect of ensuring that forensic laboratories can satisfy minimum standards.[320]

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."[321]

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."[322]

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."[323]

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."[324]

2. FORENSIC DETERMINATION OF UNCERTAINTY AND UNRELIABILITY

a. FUNDAMENTAL TO UNDERSTANDING OF TEST RESULT

---

[317] O'Dell, *A quality assurance system for DNA testing* 2(1) FOR. SCI. J. __ (2003).
[318] NAS, *Strengthening Forensic Science in the United States: A Path Forward*, 7-18 (2009); FQS-I, *Position Statement Regarding NAS Report* (2009).
[319] ASCLD/LAB – International, *Lab International Accreditation Program*, 3 (2006).
[320] SWGDRUG, *Recommendations (Minimum Standards)*, 25 (2008); Gullberg, *Results of a Proposed Breath Alcohol Proficiency Test Program*, 51(1) J. For. Sci. 168,168 (2006).
[321] FQS-I, *Proficiency Testing*, FRAP-2, 3 (2009).
[322] 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).
[323] Gullberg, *Estimating the measurement uncertainty in forensic breath-alcohol analysis*, 11 ACCRED. QUAL. ASSUR. 562, 563 (2006).
[324] Gullberg, *Methodology and Quality Assurance in Forensic Breath Alcohol Analysis*, 12 For. Sci. Rev. 49, 50 (2000).

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?"[325]

ii. "Only when measurement 'error' is acknowledged and properly estimated can…analytical goals [be] achieved."[326]

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."[327]

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."[328]

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"[329]

b. ISO/NIST METHODOLOGY: Forensic determination and reporting of uncertainty is governed by the requirements of ISO 17025[330] and NIST 1297.[331]

c. FORENSIC DETERMINATION OF UNCERTAINTY:[332]

i. BEST ESTIMATE OF MEASURAND VALUE:

a) WEIGHTED MEAN

1) "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."[333]

---

[325] 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).
[326] Gullberg, *Estimating the measurement uncertainty in forensic breath-alcohol analysis*, 11 ACCRED. QUAL. ASSUR. 562, 563 (2006).
[327] 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).
[328] NAS, *Strengthening Forensic Science in the United States: A Path Forward*, 4-9 (2009).
[329] Gullberg, *Statistical Applications in Forensic Toxicology*, Medical-Legal Aspects of Alcohol, p. 457, 458 (James Garriott ed., 5th ed. 2009).
[330] ILAC, *Guideline for Forensic Science Laboratories*, ILAC G19, § 5.10 (2002); ASCLD/LAB – *International, Estimating Uncertainty of Measurement Policy*, 1 (2007).
[331] FQS-I, *Uncertainty of Measurement*, FRAP-3, § 2.1 (2008); Ted Vosk, *The DataMaster*, Defending DUIs in Washington Ch.13 (3rd Ed. 2008).
[332] Ted W. Vosk, *Uncertainty in Forensic Breath Alcohol Testing*, Intoxication Test Evidence, Ch. 56, (2nd Ed. 2009).
[333] Gullberg, *Using a Weighted Mean to Compute the Values of Simulator Solution Standards*, 14(3) J. ANAL. TOXICOL. 196-8 (1990).

a) "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."[334]

b) "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."[335]

2) "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*."[336]

ii. UNCERTAINTY:[337]

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."[338]

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."[339]

c) Forensic labs must "[c]onstruct and document an appropriate measurement uncertainty budget, identifying and listing all potential sources of uncertainty."[340]

d) "Records must be maintained to describe the process used to develop the estimation of uncertainty. These records must include the elements of the

---

[334] Gullberg, *Using a Weighted Mean to Compute the Values of Simulator Solution Standards*, 14(3) J. ANAL. TOXICOL. 196-8 (1990).
[335] Gullberg, *Using a Weighted Mean to Compute the Values of Simulator Solution Standards*, 14(3) J. ANAL. TOXICOL. 196-8 (1990).
[336] Flanagan, *Fundamentals of Analytical Toxicology* 370 (Wiley 2007).
[337] Ted W. Vosk, *Uncertainty in Forensic Breath Alcohol Testing*, Intoxication Test Evidence, Ch. 56, (2nd Ed. 2009); Ted Vosk, *The DataMaster*, Defending DUIs in Washington Ch.13 (3rd Ed. 2008).
[338] Gullberg, *Estimating the measurement uncertainty in forensic breath-alcohol analysis*, 11 ACCRED. QUAL. ASSUR. 562, 568 (2006).
[339] Gullberg, *Methodology and Quality Assurance in Forensic Breath Alcohol Analysis*, 12 For. Sci. Rev. 49, 50 (2000).
[340] ASCLD/LAB – International, *ESTIMATING UNCERTAINTY of MEASUREMENT POLICY*, 2 (2007).

[uncertainty] budget, data gathered, calculations to arrive at the estimate, and the estimated uncertainty associated with the measurement method."[341]

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.[342]

b) The generally accepted coverage factor is $\lambda = 2$ or 3, yielding a level of confidence is 95% - 99%.[343]

$$\begin{aligned} \mathrm{Y} &= \mathrm{y} \pm U \\ &= \mathrm{y} \pm \lambda\mu_{\mathrm{c}} \end{aligned}$$

iv. THE SAFETY MARGIN

a) A "valid approach is to make a deduction for uncertainty before the final result is reported to the court."[344] 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."[345]

$$\mathrm{Y} \geq y - s$$

d. FORENSIC DETERMINATION OF UNRELIABILITY

i. "Understanding the variables inherent in a measurement system that deals with [qualitative characteristics] is fundamental to a good quality assurance program."[346]

ii. FREQUENTIST METHODS: When engaging in qualitative analysis, "the paradigm of yes/no conclusions is useful for describing and quantifying the accuracy with which

[341] ASCLD/LAB – International, *ESTIMATING UNCERTAINTY of MEASUREMENT POLICY*, 3 (2007).
[342] 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).
[343] 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).
[344] 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).
[345] Treble, *Analytical measurement and the law* 20 VAM BULLETIN 3, 4 (1999).
[346] Reeder, *Impact of DNA Typing on Standards and Practice in the Forensic Community* 123 ARCH. PATH. LAB. MED. 1063, 1064 (1999).

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."[347]

	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.[348]

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."[349]

iii. BAYESIAN METHODS: "Bayesian estimates are particularly widely used in evaluating forensic evidence, for example DNA matching or blood group matching."[350]

a) LIKELIHOOD RATIOS: The likelihood ratio "approach is now firmly established as a theoretical framework for any forensic discipline."[351] Since the likelihood ratio is a measure of the impact of the evidence on the initial hypothesis H, in a

---

347 NAS, *Strengthening Forensic Science in the United States: A Path Forward*, 4-6 (2009).
348 NAS, *Strengthening Forensic Science in the United States: A Path Forward*, 4-7 (2009).
349 NAS, *Strengthening Forensic Science in the United States: A Path Forward*, 4-8 – 4-9 (2009).
350 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).
351 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, *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).

prosecution that the defendant is guilty/not guilty, it provides a quantitative measure of the relevance and weight of the evidence.[352]

b) <u>POSTERIOR PROBABILITY</u>:[353]

c) <u>BAYESIAN NETWORKS</u>: "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."[354]

---

[352] Stockmarr, *Likelihood Ratios for Evaluating DNA Evidence When the Suspect is Found Through a Database Search* 55 BIOMETRICS 671 (1999).
[353] Meester, *Why the Effect of Prior Odds Should Accompany the Likelihood Ratio When Reporting DNA Evidence* 3 LAW, PROB. AND RISK 51 (2004).
[354] 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).

3. REPORTING FORENSIC RESULTS[355]

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."[356]

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 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."[357]

iii. "The terminology used in reporting and testifying about the results of forensic science investigations must be standardized."[358]

iv. "Calculations and data transfers which do not form part of a validated electronic process should be checked, preferably by a second person."[359]

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."[360]

---

[355] Ted W. Vosk, *Uncertainty in Forensic Breath Alcohol Testing*, Intoxication Test Evidence, Ch. 56, (2nd Ed. 2009); Ted Vosk, *The DataMaster*, Defending DUIs in Washington Ch.13 (3rd Ed. 2008).
[356] Gullberg, *Professional and Ethical Considerations in Forensic Breath Alcohol Testing Programs* 5(1) J. ALC. TEST. ALLIANCE 22, 25 (2006).
[357] Gullberg, *Statistical Applications in Forensic Toxicology*, Medical-Legal Aspects of Alcohol, p. 457, 504 (James Garriott ed., 5th ed. 2009).
[358] NAS, *Strengthening Forensic Science in the United States: A Path Forward*, S-15 (2009).
[359] ILAC, *Guidelines for Forensic Science Laboratories*, ILAC-G19 § 4.12.2.1(e) (2002).
[360] SOFT/AAFS, *Forensic Toxicology Laboratory Guidelines*, § 10.1 (2006).

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."[361]

a) If an outlier is suspected then it can be investigated utilizing Grubb's test.[362]

b. RESULT = MEASUREMENT + UNCERTAINTY

i. "[F]orensic test results must be validated and verified before they are presented to the court."[363]

ii. "If systematic error does exist this must be added or subtracted from the mean result."[364]

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." [365] "Estimating and reporting measurement uncertainty with the number completes the picture and allows us to properly use the result to make reliable and defensible decisions."[366]

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."[367]

v. "All results for every forensic science method should indicate the uncertainty in the measurements that are made."[368]

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."[369]

---

361 SOFT/AAFS, *Forensic Toxicology Laboratory Guidelines*, § 8.3.9 (2006).
362 Flanagan, *Fundamentals of Analytical Toxicology* 385 (Wiley 2007).
363 Godowsky, *Quality Assurance in forensic Laboratories*, 4(6) EV. TECH. MAG. 36, 36 (2006).
364 Jones, *Dealing with Uncertainty in Chemical Measurements*, 14(1) NEWSL. INT. ASSOC. CHEM. TEST. 10 (2003).
365 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).
366 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).
367 Gullberg, *Statistical Applications in Forensic Toxicology*, Medical-Legal Aspects of Alcohol, p. 457, 504-505 (James Garriott ed., 5th ed. 2009).
368 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).
369 NAS, *Strengthening Forensic Science in the United States: A Path Forward*, S-16, 6-3 (2009).

vii. "[C]onsidering or not the uncertainty of a critical result can make the difference between acquittal and a guilty sentence."[370]

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."[371]

1) "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."[372]

2) "An urgent need exists to report results of forensic alcohol analysis as a range of values, that is as a confidence statement."[373]

b) Forensic results need to be reported, along with a coverage interval that has a high probability of containing the true value of the measurand, that is "the mean plus or minus two standard deviations."[374]

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.[375]

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."[376]

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

[370] Bich, *Interdependence between measurement uncertainty and metrological traceability* ACCRED. QUAL. ASSUR. (IN PRESS - 2009).
[371] Gullberg, *Estimating the measurement uncertainty in forensic breath-alcohol analysis*, 11 ACCRED. QUAL. ASSUR. 562, 568 (2006).
[372] 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).
[373] A.W. Jones, Ph.D, *Dealing with Uncertainty in Chemical Measurements*, 14(1) Newsletter of the International Association for Chemical Testing 6, 7 (2003).
[374] NAS, *Strengthening Forensic Science in the United States: A Path Forward*, 4-6 (2009).
[375] 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).
[376] Taroni, *Bayesian Networks and Probabilistic Inference in Forensic Science*, Preface (Wiley 2006).

relative strength and impact of the evidence on the determination to be made so that it may be assigned appropriate weight.[377]

F. Scientific Forensic Standards[378]

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."[379]

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."[380]

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."[381]

d. Standardization in forensic science reduces confusion, eliminates causes of error and makes it possible for independent evaluation of results.[382]

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."[383]

2. Gold Standards: Standards promulgated by ISO, NIST and ASTM are nearly universally recognized throughout the forensics community.[384]

a. ISO:

[377] Champod, *Identification and Individualization* in *Wiley Encyclopedia of Forensic Science* (*in press* 2009).
[378] Ted Vosk, *The DataMaster*, Defending DUIs in Washington Ch.13 (3rd Ed. 2008).
[379] NAS, *Strengthening Forensic Science in the United States: A Path Forward*, 7-18 (2009).
[380] Lentini, *Forensic Science Standards: Where They Come From and How They Are Used*, 1 For. Sci. Pol. Mgmt. 10, 16 (2009).
[381] Lentini, *Forensic Science Standards: Where They Come From and How They Are Used*, 1 For. Sci. Pol. Mgmt. 10, 16 (2009).
[382] 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).
[383] FQS-I, *Position Statement Regarding NAS Report* (2009).
[384] 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); Ted Vosk, *The DataMaster*, Defending DUIs in Washington Ch.13 (3rd Ed. 2008); 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).

i. "Any laboratory seeking ASCLD/LAB–International accreditation must demonstrate conformance to the requirements in ISO/IEC 17025."[385] "Conforming to the numbered requirements in [ISO 17025] is mandatory."[386]

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."[387]

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."[388]

3. "Appropriate [forensic] standards must be coupled with effective systems of accreditation and/or certification that include strong enforcement mechanisms and sanctions."[389]

VI. Special Topics

A. DNA

1. Statistical Approach to Presentation of Results[390]

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."[391]

b. Research Findings – what to look out for: [392]

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."

---

[385] ASCLD/LAB – International, *Lab International Accreditation Program*, 2 (2006); ASCLD/LAB – International, *Breath Alcohol Calibration Accreditation Program*, 3 (2008).
[386] ASCLD/LAB – International, *Lab International Accreditation Program*, 2 (2006).
[387] NIST, *Office of Law Enforcement Standards: Programs, Activities and Accomplishments*, NISTIR 7366, 26 (2007); Ted Vosk, *The DataMaster*, Defending DUIs in Washington Ch.13 (3rd Ed. 2008).
[388] Lentini, *Forensic Science Standards: Where They Come From and How They Are Used*, 1 For. Sci. Pol. Mgmt. 10, 12-15 (2009).
[389] NAS, *Strengthening Forensic Science in the United States: A Path Forward*, 7-10 (2009).
[390] Koehler, *When Are People Persuaded By DNA Match Statistics?* 25(5) Law and Human Behavior 493 (2001).
[391] Koehler, *When Are People Persuaded By DNA Match Statistics?* 25(5) Law and Human Behavior 493, 493 (2001).
[392] Koehler, *When Are People Persuaded By DNA Match Statistics?* 25(5) Law and Human Behavior 493, 494-495 (2001).

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:[393]

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:[394]

a) Recognizing possibility that DNA match arose by coincidence.

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

b) Realizing that coincidental matches exist and are plentiful.

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

d. Strategic and policy implications:[395]

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

---

393 Koehler, *When Are People Persuaded By DNA Match Statistics?* 25(5) LAW AND HUMAN BEHAVIOR 493, 495 (2001).
394 Koehler, *When Are People Persuaded By DNA Match Statistics?* 25(5) LAW AND HUMAN BEHAVIOR 493, 508-509 (2001).
395 Koehler, *When Are People Persuaded By DNA Match Statistics?* 25(5) LAW AND HUMAN BEHAVIOR 493, 509 (2001).

error need to be explored and quantified in order to arrive at reliable error rate estimates for DNA analysis."[396]

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."[397]

i. Sources of false positives include:[398]

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."[399]

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."[400]

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.[401]

---

396 NAS, *Strengthening Forensic Science in the United States: A Path Forward*, 4-8 – 4-9 (2009).
397 Thompson, *How the Probability of a False Positive Affects the Value of DNA Evidence* 48(1) J. For. Sci. 47 (2003).
398 Thompson, *How the Probability of a False Positive Affects the Value of DNA Evidence* 48(1) J. For. Sci. 47 (2003).
399 Thompson, *How the Probability of a False Positive Affects the Value of DNA Evidence* 48(1) J. For. Sci. 47, 56 (2003).
400 NCCLS, *Blood Alcohol Testing in the Clinical Laboratory; Approved Guideline*, T/DM6-A § 1.4.6 (1997).
401 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

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.”[402]

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.”[403]

ii. “Aqueous Standard Reference Materials containing ethanol are available from the National Institute of Standards and Technology (SRM 1828a).”[404]

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.”[405]

ii. “Every analysis or batch of analyses should be accompanied by the analysis of negative and positive controls.”[406]

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.”[407]

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

---

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).

402 NCCLS, *Blood Alcohol Testing in the Clinical Laboratory; Approved Guideline*, T/DM6-A § 5 (1997).

403 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).

404 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).

405 NCCLS, *Blood Alcohol Testing in the Clinical Laboratory; Approved Guideline*, T/DM6-A § 5.1 (1997).

406 NCCLS, *Blood Alcohol Testing in the Clinical Laboratory; Approved Guideline*, T/DM6-A § 5.2 (1997).

407 NCCLS, *Blood Alcohol Testing in the Clinical Laboratory; Approved Guideline*, T/DM6-A § 5.2 (1997).

8105 NE 140th Pl., Bothell WA 98011

one positive and one negative control. For qualitative assays positive and negative controls, acceptable results may simply be positive or negative, respectively."[408]

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."[409]

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.[410]

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."[411]

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."[412] "The errors that can occur during the collection and handling of blood specimens are potentially numerous"[413]

a) "Nonbiological factors—such as patient misidentification…contribute to the total 'laboratory error.'"[414]

[408] SOFT/AAFS, *Forensic Toxicology Laboratory Guidelines*, § 9.2.1 (2006).
[409] SOFT/AAFS, *Forensic Toxicology Laboratory Guidelines*, § 9.1.9 (2006).
[410] Falkensson, *Hospital alcohol analyses not completely reliable External quality control needed at least every year* 87 LÄKARTIDNINGEN 470 (1990).
[411] NCCLS, *Procedures for the Collection of Diagnostic Blood Specimens by Venipuncture; Approved Standard—Fifth Edition*, H3-A5, vii (2003).
[412] NCCLS, *Procedures for the Collection of Diagnostic Blood Specimens by Venipuncture; Approved Standard—Fifth Edition*, H3-A5, § 5 (2003).
[413] 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).
[414] NCCLS, *Procedures for the Collection of Diagnostic Blood Specimens by Venipuncture; Approved Standard—Fifth Edition*, H3-A5, § 5 (2003).

b) "[B]iological factors—such as patient posture and the time a specimen is drawn, [] contribute to the total 'laboratory error.'"[415]

c) Contamination

1) "If the venipuncture proves difficult and the vein must be touched again to draw blood, the site should be cleansed again."[416]

2) "When drawing a blood specimen for alcohol testing, a nonalcohol-based cleanser should be used to cleanse the venipuncture site."[417]

d) <u>Hemolysis</u>: "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."[418]

1) "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."[419]

ii. <u>Specimen Handling and Storage effects</u>: "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."[420]

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."[421] The types of effects can be categorized as:[422]

---

415 NCCLS, *Procedures for the Collection of Diagnostic Blood Specimens by Venipuncture; Approved Standard—Fifth Edition*, H3-A5, § 5 (2003).
416 NCCLS, *Procedures for the Collection of Diagnostic Blood Specimens by Venipuncture; Approved Standard—Fifth Edition*, H3-A5, § 8.8.3 (2003).
417 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).
418 *Stedman's Medical Dictionary for the Health Professions and Nursing* 701 (6th ed. 2008).
419 NCCLS, *Procedures for the Collection of Diagnostic Blood Specimens by Venipuncture; Approved Standard—Fifth Edition*, H3-A5, § 10.3 (2003).
420 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).
421 Williams, *How do storage conditions affect your samples?* 20 VAM BULLETIN 22, 22 (1999).
422 Williams, *How do storage conditions affect your samples?* 20 VAM BULLETIN 22, 22 (1999).

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

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

3) 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."[423]

1) "The time of collection is critical information which must be recorded and should appear on the report of results."[424]

2) "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)."[425]

3) "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."[426]

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

[423] 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).
[424] NCCLS, *Blood Alcohol Testing in the Clinical Laboratory; Approved Guideline*, T/DM6-A § 2.3.1 (1997).
[425] 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).
[426] NCCLS, *Blood Alcohol Testing in the Clinical Laboratory; Approved Guideline*, T/DM6-A § 2.3.1 (1997).

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."[427]

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 serum/whole blood alcohol concentrations, with typical experimental ranges of 1.05/1 to 1.25/1."[428]

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."[429]

b. AN UNCERTAINTY BUDGET FOR THE MEASUREMENT OF ETHANOL IN BLOOD BY HEADSPACE GAS CHROMATOGRAPHY:[430]

i. Includes four sources of uncertainty:[431]

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.[432]

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."[433]

iii. Conclusion:

[427] Williams, *How do storage conditions affect your samples?* 20 VAM BULLETIN 22, 24 (1999).
[428] NCCLS, *Blood Alcohol Testing in the Clinical Laboratory; Approved Guideline*, T/DM6-A § 2.2 (1997).
[429] NCCLS, *Procedures for the Collection of Diagnostic Blood Specimens by Venipuncture; Approved Standard—Fifth Edition*, H3-A5, § 5 (2003).
[430] Kristiansen, *An Uncertainty Budget for the Measurement of Ethanol in Blood by Headspace Gas Chromatography*, 28(6) J. ANAL. TOX. 456 (2004).
[431] Kristiansen, *An Uncertainty Budget for the Measurement of Ethanol in Blood by Headspace Gas Chromatography*, 28(6) J. ANAL. TOX. 456, 457 (2004).
[432] Kristiansen, *An Uncertainty Budget for the Measurement of Ethanol in Blood by Headspace Gas Chromatography*, 28(6) J. ANAL. TOX. 456, 462 (2004).
[433] Kristiansen, *An Uncertainty Budget for the Measurement of Ethanol in Blood by Headspace Gas Chromatography*, 28(6) J. ANAL. TOX. 456, 462 (2004).

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."[434]

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."[435]

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."[436]

d) Conversion:[437] g/Kg → g/100ml

   1) BAC g/100mL ≡ .1055 · BAC g/Kg

   2) Safety Margin = .1 g/Kg
      = .0106 g/100ml

iv. <u>EXTENSION</u>: Assume the result compensated for by reduction of the safety margin determined is represented by $BAC \geq R_{esult} - U$ (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 $\varepsilon > U$.

---

434 Kristiansen, *An Uncertainty Budget for the Measurement of Ethanol in Blood by Headspace Gas Chromatography*, 28(6) J. ANAL. TOX. 456, 463 (2004).
435 Kristiansen, *An Uncertainty Budget for the Measurement of Ethanol in Blood by Headspace Gas Chromatography*, 28(6) J. ANAL. TOX. 456, 463 (2004).
436 Kristiansen, *An Uncertainty Budget for the Measurement of Ethanol in Blood by Headspace Gas Chromatography*, 28(6) J. ANAL. TOX. 456, 463 (2004).
437 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. <u>Different Concepts Employed as Concentration</u>: [438]

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.[439] Reporting conventions include:[440]

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.[441] The metrologically sound practice would be to standardize units and computations so that results reported by different individuals could be readily compared and understood.[442]

[438] 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).

[439] 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).

[440] 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).

[441] 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).

[442] 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).

8105 NE 140th Pl., Bothell WA 98011

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."[443]

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."[444]

3. "[I]n order to qualify as 'scientific knowledge,' an inference or assertion must be derived by the scientific method."[445]

B. DAUBERT,[446] FRYE[447] AND EVID. R. 702:

1. "In a case involving scientific evidence, *evidentiary reliability* will be based upon *scientific validity*."[448] "The term 'scientific' implies a grounding in the methods and procedures of science."[449] It "draws its convincing force from some principle of science, mathematics and the like."[450] "[A] hypothesis that cannot be subject to the possibility of rejection by observation and experiment cannot be regarded as scientific."[451] "[I]ndeed this methodology distinguishes science from other fields of human inquiry."[452]

2. Under *Daubert*, courts must engage in "a preliminary assessment of whether the reasoning or methodology underlying the testimony is scientifically valid."[453] 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

---

443 Justice Stephen Breyer in, *Reference Manual on Scientific Evidence* 4 – 8 (2nd ed. 2000).
444 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").
445 *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").
446 *Daubert v. Merrell Dow Pharmaceuticals, Inc.*, 509 U.S. 579, 590 (1993).
447 *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.
448 *Daubert v. Merrell Dow Pharmaceuticals, Inc.*, 509 U.S. 579, 590 n.9 (1993).
449 *Reese v. Stroh*, 874 P.2d 200, 206 (1994).
450 *State v. Brown*, 687 P.2d 751, 754 (Or. 1984).
451 *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).
452 *Daubert v. Merrell Dow Pharmaceuticals, Inc.*, 509 U.S. 579, 593 (1993).
453 *Daubert v. Merrell Dow Pharmaceuticals, Inc.*, 509 U.S. 579, 592-593 (1993).

8105 NE 140th Pl., Bothell WA 98011

technology may rest on past accomplishments nor may the courts."[454] "The focus, of course, must be solely on principles and methodology, not on the conclusions that they generate."[455] The following factors are relevant to the determination of scientific reliability.

3. VALIDITY[456]

a. <u>Daubert v. Merrell Dow Pharmaceuticals, Inc.</u>, 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. <u>State v. Lasworth</u>, 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

[454] *U.S. v Horn*, 185 F.Supp.2d 530, 536 Fn.15 (D.Md. 2002).
[455] *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.").
[456] *Daubert v. Merrell Dow Pharmaceuticals, Inc.*, 509 U.S. 579, 591 (1993).

BACs at or 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[457]

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)[458]

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

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

which their calculations relied. The testimony also established that the method of calculation 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. <u>*DeLuca by DeLuca v. Merrell Dow Pharmaceuticals, Inc.*</u>, 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>*U.S. v. Downing*</u>, 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. <u>*Thomas v. Allen*</u>, 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."

8105 NE 140th Pl., Bothell WA 98011

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. <u>*Henricksen v. ConocoPhillips Co.*</u>, __ 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).[459]

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).

[459] 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.

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 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. <u>*State v. Brown*</u>, 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. <u>*Com. v. Curnin*</u>, 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. <u>*State v. Keller*</u>, 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. <u>*State v. Boehmer*</u>, 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[460]

   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.

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

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 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. <u>Milanowicz v. The Raymond Corp.</u>, 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. <u>Ex parte Taylor</u>, 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. <u>People v. Shreck</u>, 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[461]

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

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

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, *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."

8105 NE 140th Pl., Bothell WA 98011

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."[462]

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.

[462] NIST, OLES Mission Statement http://www.eeel.nist.gov/oles/oles_mission.html.

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."[463]

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."[464]

b) "The general requirements in sections 4 and 5 incorporate ISO/IEC 17025:2005 (as adopted by the NVLAP Calibration Laboratories Accreditation Program) and address internationally accepted quality management practices for calibration and testing laboratories."[465]

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

*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

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).

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.

463 NIST, *State Weights and Measures Laboratories, Program Handbook*, NIST HB 143, 11 (2007).
464 NIST, *State Weights and Measures Laboratories, Program Handbook*, NIST HB 143, 11 (2007).
465 NIST, *State Weights and Measures Laboratories, Program Handbook*, NIST HB 143, 11 (2007).

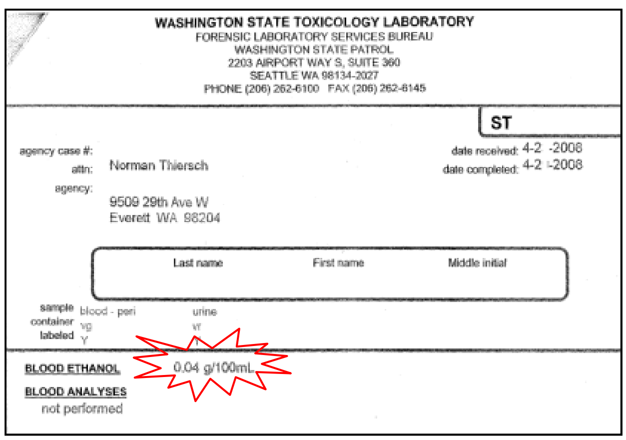
b. *United States v. Van Griffin*, 874 F.2d 634, 638 (9th Cir. 1989)(Standards issued by agency responsible for subject matter is evidence standards are necessary for reliability).

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

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

1. FAILURE TO REPORT UNCERTAINTY WITH BLOOD TEST RESULTS[466]


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: blood - peri / urine
container: vg / vr
labeled: Y / Y

BLOOD ETHANOL 0.04 g/100mL

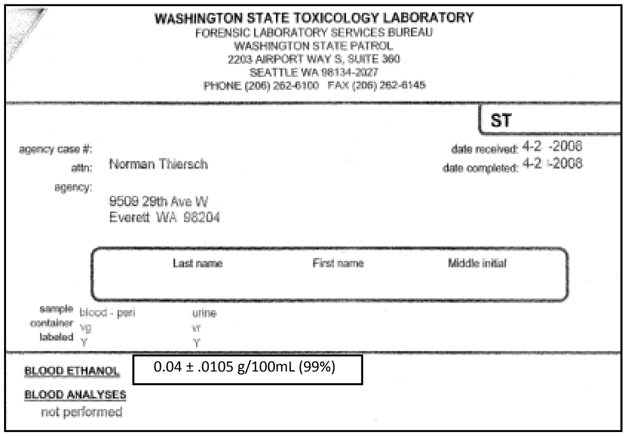
BLOOD ANALYSES
not performed

---

[466] Ted W. Vosk, *Uncertainty in Forensic Breath Alcohol Testing*, Intoxication Test Evidence, Ch. 56, (2nd Ed. 2009).

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

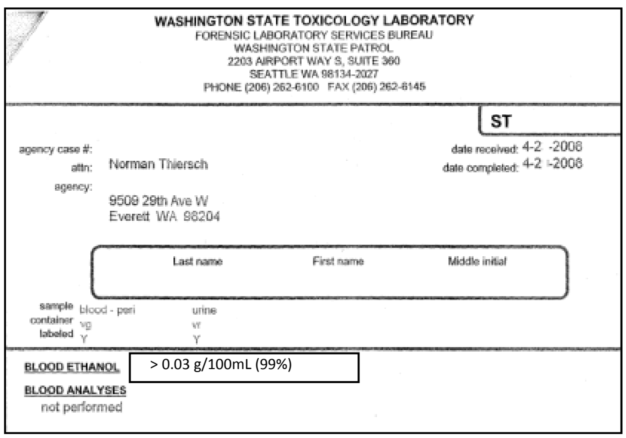
date received: 4-2 -2008
date completed: 4-2 -2008

Last name First name Middle initial

sample container labeled: blood - peri, vg, Y; urine, vr, Y

BLOOD ETHANOL 0.04 ± .0105 g/100mL (99%)
BLOOD ANALYSES
not performed

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, vg, Y; urine, vr, Y

BLOOD ETHANOL > 0.03 g/100mL (99%)
BLOOD ANALYSES
not performed

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

1. <u>WAC 448-14-020</u>: 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.

IX. CASE STUDY: WASHINGTON STATE TOXICOLOGY LAB – *State v Ahmach*, King County District Court (Jan. 2008).[467]

A. VIOLATED STANDARDS – EXAMPLES

1. TRACEABILITY: "There was no documentation available to show what reagents and controls were used in the testing and therefore traceability of the reagent used to prepare these certified solutions cannot be documented…It was reported that the material used to prepare the solutions was purchased from the '*liquor store*'. Did it come with a certificate of authenticity or traceability? Documentation needs to be maintained as to the source and quality of these reagents." ASCLD Audit. ISO 17025 § 5.10.3 (2005); NIST, SOP-1, § 2 (2003); *GUM* § 7.1.4 (2008); SWGDRUG, *Recommendations (Minimum Standards)*, 31 (2008).

National Institute of Standards and Technology

Certificate of Analysis

Standard Reference Material® 1847

Ethanol-Water Solutions (Breath-Alcohol Testing: Three Levels)

This Standard Reference Material (SRM) is a set of three levels of ethanol (ethyl alcohol: Chemical Abstracts Service [CAS] Registry Number 64-17-5) in water. SRM 1847 is intended primarily for use in the calibration of instruments and techniques used for the determination of ethanol in breath. The three levels (nominal concentrations of 2 %, 6 %, and 25 % by mass) are available as a set in SRM 1847 and as individual solutions in SRMs 2897 through 2899, respectively. A second SRM suite, SRM 1828b, is intended primarily for the calibration of instruments and techniques used for the determination of ethanol in blood. SRM 1828b is a set of six levels (nominal concentrations of 0.02 %, 0.04 %, 0.08 %, 0.1 %, 0.2 %, and 0.3 % by mass). These six levels are also available as individual solutions in SRMs 2891 through 2896, respectively. A unit of SRM 1847 consists of six 10-milliliter ampoules (two ampoules each of the nominal concentrations of 2 %, 6 %, and 25 % by mass), each containing approximately 10 mL of solution.

[467] See Appendix B for decision; See also, Ted Vosk, *Chaos Reigning: Breath Testing and the Washington State Toxicology Lab*, THE NACDL CHAMPION 54 (May/June 2008); Ted Vosk, *The DataMaster*, Defending DUIs in Washington Ch.13 (3rd Ed. 2008).

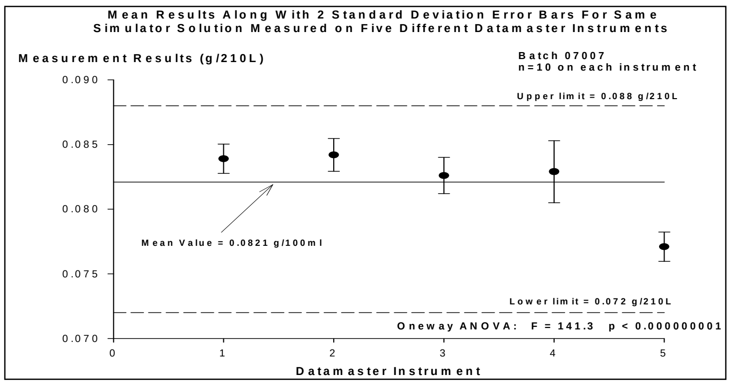
2. FAILURE TO VALIDATE SOFTWARE: ISO 17025 § 5.4.7.2 (2005); NIST HB 150 § 5.4.7.2 (2006).

Washington State Patrol Toxicology Laboratory
2203 Airport Way S.,
Seattle WA 98004

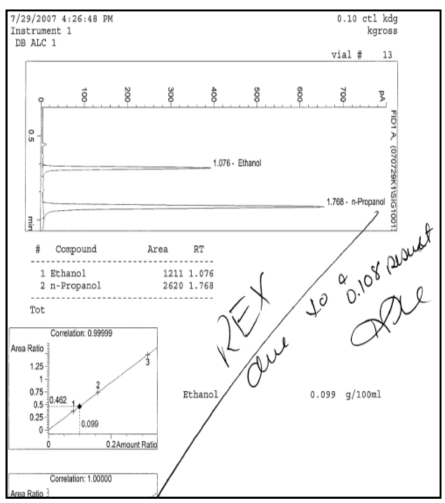
August 9, 2007

Calculation Error in Program used to calculate Reference Values for External Standard Simulator Solutions.

3. FAILURE TO ACCOUNT FOR BIAS: *(GUM)*, § 6.3.2 (2008); NIST TN 1297, §5.2, App. D (1994).

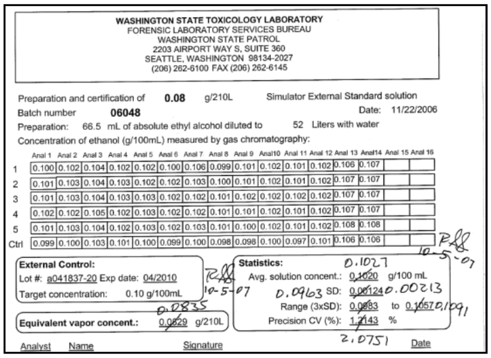


4. OUTLIERS – DISCARDING VALID DATA: ASTM E 178 (2008); NIST SP260-100 (1993).



8105 NE 140th Pl., Bothell WA 98011

5. Failure to Check Data or Calculations: ISO 17025 § 5.4.7.1 (2005); NIST, HB 150 § 5.4.7.1 (2006).



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

Preparation and certification of **0.08** g/210L Simulator External Standard solution
Batch number **06048** Date: 11/22/2006
Preparation: 66.5 mL of absolute ethyl alcohol diluted to 52 Liters with water
Concentration of ethanol (g/100mL) measured by gas chromatography:

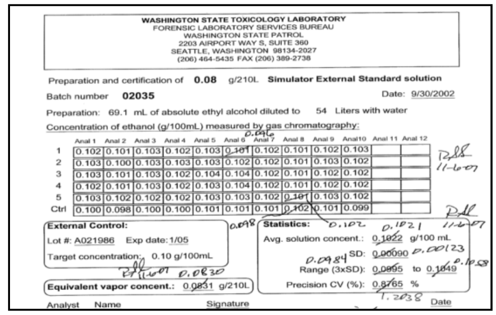
	Anal 1	Anal 2	Anal 3	Anal 4	Anal 5	Anal 6	Anal 7	Anal 8	Anal 9	Anal10	Anal 11	Anal 12	Anal 13	Anal14	Anal 15	Anal 16
1	0.100	0.102	0.104	0.102	0.102	0.100	0.106	0.099	0.101	0.102	0.101	0.102	0.106	0.107		
2	0.101	0.103	0.104	0.103	0.102	0.102	0.103	0.100	0.101	0.102	0.101	0.102	0.107	0.107		
3	0.101	0.103	0.104	0.103	0.103	0.102	0.102	0.101	0.102	0.102	0.102	0.102	0.107	0.107		
4	0.102	0.102	0.105	0.102	0.103	0.102	0.102	0.101	0.101	0.102	0.101	0.102	0.107	0.107		
5	0.101	0.103	0.104	0.102	0.103	0.102	0.103	0.101	0.100	0.102	0.101	0.102	0.108	0.108		
Ctrl	0.099	0.100	0.103	0.101	0.100	0.099	0.100	0.098	0.098	0.100	0.097	0.101	0.106	0.106		

External Control:
Lot #: a041837-20 Exp date: 04/2010
Target concentration: 0.10 g/100mL

Equivalent vapor concent.: 0.0829 g/210L

Statistics:
Avg. solution concent.: 0.1020 g/100 mL
SD: 0.00124
Range (3xSD): 0.0983 to 0.1057
Precision CV (%): 1.2143 %

Analyst Name Signature Date



WASHINGTON STATE TOXICOLOGY LABORATORY
FORENSIC LABORATORY SERVICES BUREAU
WASHINGTON STATE PATROL
2203 AIRPORT WAY S, SUITE 360
SEATTLE, WASHINGTON 98134-2027
(206) 464-5435 FAX (206) 389-2738

Preparation and certification of **0.08** g/210L **Simulator External Standard solution**
Batch number **02035** Date: 9/30/2002
Preparation: 69.1 mL of absolute ethyl alcohol diluted to 54 Liters with water
Concentration of ethanol (g/100mL) measured by gas chromatography:

	Anal 1	Anal 2	Anal 3	Anal 4	Anal 5	Anal 6	Anal 7	Anal 8	Anal 9	Anal10	Anal 11	Anal 12
1	0.102	0.101	0.103	0.102	0.103	0.101	0.102	0.101	0.102	0.103		
2	0.103	0.100	0.103	0.103	0.103	0.102	0.102	0.101	0.103	0.102		
3	0.103	0.101	0.103	0.102	0.104	0.104	0.102	0.101	0.102	0.102		
4	0.102	0.101	0.103	0.102	0.103	0.104	0.102	0.101	0.102	0.102		
5	0.103	0.102	0.102	0.102	0.103	0.103	0.102	0.101	0.103	0.102		
Ctrl	0.100	0.098	0.100	0.100	0.101	0.101	0.101	0.102	0.101	0.099		

External Control:
Lot #: A021986 Exp date: 1/05
Target concentration: 0.10 g/100mL

Equivalent vapor concent.: 0.0831 g/210L

Statistics:
Avg. solution concent.: 0.1022 g/100 mL
SD: 0.00090
Range (3xSD): 0.0995 to 0.1049
Precision CV (%): 0.8765 %

Analyst Name Signature Date

B. Evidentiary Ruling[468]

1. Uncertainty and Standards

a. "…even errors in the range of 1 or 2% can have a profound effect on a breath test reading. Nonetheless, each expert witness who offered testimony stated that there was not a process or a machine that would not insert some amount of inherent error in any result…A breath test machine normally has a bias of 1-2%, with the smaller fraction of the machines registering a bias of 5% or less. The breath test program is not, however, ser up to account for any of the potential bias inherent in the breath test machine. Thus, a process that already allows potential bias in each reading only underscores the importance of ensuring that the WSTL eliminates all other possible sources of error."

b. "Dr. Nayak Polissar, an expert called by the State, testified that only superior methods will ensure accuracy, and that the accuracy and precision necessary for a particular laboratory task is dependent upon the particular use intended for the final product. As stated by the National Institute of Standards and Technology (NIST), 'accuracy…is judged with respect to the use to be made of the data.' NIST Special Publication 260-100, 2 (1993)."

2. Violations of Recognized Standards

a. "The WSTL did not require that the data transfer be checked, and toxicologists signed certifications which were unverified and later found incorrect. Many errors in diverse areas were subsequently discovered."

b. "The computer software used to enter and calculate…lab results on the worksheets was not created by an individual with the requisite knowledge and skill necessary to ensure that the data was correctly analyzed and recorded. Moreover, no one checked the software to determine if it was operating properly…The WSTL itself never considered that it was necessary to check the software to ensure that it was fit for purpose. The software contained errors which were not revealed until the WSTL came under close scrutiny because of the [perjury] investigation."

468 See Appendix B for Evidentiary Ruling.

c. "Literally thousands of breath tests performed in recent years were affected through a multiplicity of errors in the toxicology lab. A very brief recitation of the errors include: the improper rejection of data; erroneously switched data…the use of software that improperly computed data and that improperly ignored [] data; and, the use of simulator solutions [reference materials] that were outside of the allowable range."

3. EVID. R. 702

a. "[U]nder ER 702, the work product of the WSTL is sufficiently compromised by ethical lapses, systemic inaccuracy, negligence and violations of scientific principles that the WSTL simulator solution work product would not be helpful to the trier of fact…without reliable evidence that a correctly functioning breath test instrument can provide, the discovery of truth in DUI cases suffers; the innocent may be wrongly convicted, and the guilty may go free."

i. "The WSTL must establish procedures that, in the years ahead, ensure that their processes are double checked for accuracy."

4. STANDARDS, ACCREDITATION AND SCIENTIFIC RELIABILITY

a. "…the Supreme Court agreed with the statement that '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'…the proposition that robust scientific standards are expected in the WSTL still remains…In the summer of 2008 the WSTL plans to adopt the General Requirements for the Competence of Testing and Calibration Laboratories, ISO/IEC 17025:1999(E), promulgated by the International Organization for Standardization. These standards are neither required for a toxicology laboratory, nor are they a panacea for the past and current problems in the WSTL. Their adoption, however, is likely to move the WSTL a long way toward the type of reliable forensic science which should be expected of a state toxicology lab."

5. RULING

a. "We hold that, under ER 702, the work product of the WSTL has been so compromised by ethical lapses, systemic inaccuracy, negligence and violations of scientific principles that the WSTL simulator solution work product would not be helpful to the trier of fact."

X. RECENT DECISION INCORPORATING NAS REPORT: *State v Berry*, Pierce County District Court (Dec. 2009).[469]

A. GENERAL SCIENTIFIC PRINCIPLES AND THE LAW

1. "While forensic science is distinct from research science some may believe that a lesser standard is acceptable. Such a conclusion would be erroneous…Adherence to scientific principles is important for concrete reasons: they enable the reliable inference of knowledge

[469] See Appendix C for decision.

from uncertain information – exactly the challenge faced by forensic scientists. Thus, the reliability of forensic science methods is greatly enhanced when those principles are followed."

2. Statute, precedent and the rules of evidence "share a universal requirement – that the underlying science and methodology supporting the evidence, in this case a measurement, must be accepted by the scientific community as reliable and that the measurement is accurate and helpful to the trier of fact."

3. "When evidence of measurement relies on inadequate scientific foundation, the proffered evidence must be classified as untrustworthy and inadmissible. To admit bogus and misleading science under the pretext of legitimate science is irrational and harmful to any notion of justice."

4. "Properly measuring and employing tested methodology prevents the introduction of bogus science. Forensic evidence that conflicts with applicable scientific principles is too unreliable to be considered by the court. This evidence is not admissible."

B. UNCERTAINTY AND REFERENCE MATERIALS

1. "In any case…uncertainty must be accounted for."

2. "Failing to consider…uncertainty is important because the simulator solution produced by the Toxicology Lab is the only means to calibrate the datamaster."

3. "In the instant case the omission of calculating for…uncertainty permeates the development and use of the reference solutions…Due to…failing to adjust for uncertainty, the concentration of any reference solution is not scientifically reliable or accurate."

4. "…the omission of outliers, bias and uncertainty, when measuring reference solutions alone is sufficient to conclude that a professional scientist would not rely on the work product of the Toxicology Lab."

C. RULING

1. "The breath measurement is not grounded on a firm scientific foundation and cannot be considered to be reliable, accurate or helpful to the trier of fact. The Toxicology Lab has not demonstrated adherence to accepted scientific principles, consequently the breath measurement is inaccurate and inadmissible."

2. "Defendant's motion to suppress is granted. This order will remain in effect until the Toxicology Lab methodology and operation is certified to ISO standards."

XI. Exercises

A. What is the importance of scientific standards:

1. Under ER 702?

2. Scientifically?

B. Why is it necessary to report measurement uncertainty with every result?

C. What standard constitutes the Gold Standard and basis for most laboratory accreditation schemes worldwide?

D. If the result of a measurement is not traceable to a recognized standard, what does the value reported mean?

E. When presenting evidence to a jury, why might reporting a likelihood ratio for qualitative analysis be more appropriate than reporting a categorical conclusion?

1. What if the categorical conclusion is accompanied by a frequentist measure of reliability?

F. Assume a measurement instrument has been calibrated at quantity values of .04, .08, .10 and .15.

1. What can be said of a measurement yielding a value of .20?

2. What can be said of a measurement yielding a value of .02?

3. Should the result be admissible under ER 702?

   a. If so, should there be any limits put on the manner in which it is reported to the jury?

G. Duplicate analysis of blood by chromatograph yields results of y_1 and y_2.

1. What information is needed in order to interpret the result?

2. Assuming the best estimate can be based on a bias adjusted mean of $\bar{y} = \frac{y_1 + y_2}{2}$, with expanded uncertainty U, how should the results be reported?

3. What is the individual's actual BAC?

H. Assume new research incorporating specimen collection and handling components of uncertainty reports an expanded uncertainty ($k = 3$) in the analysis of blood by gas chromatography of .15 mg/g.

1. Assuming your jurisdiction reports blood alcohol concentration in units of g/100mL, how much of a safety margin are you going to subtract from the result for the jury?

I. Suppose a forensic scientist reports a match between DNA samples. What are the two levels of uncertainty that must be accounted for?

8105 NE 140th Pl., Bothell WA 98011

J. Are genetic population statistics based on frequentist or Bayesian methodology?

K. In reporting the reliability of a qualitative method, when should one rely on frequentist or Bayesian methods?

1. From a scientific standpoint?

2. From a legal standpoint?

L. What is the purpose of accreditation?

M. Who is responsible for ensuring that forensic evidence presented to the court is analyzed critically?

N. Suppose you have the following table characterizing the results of validation testing of a qualitative method.

	Test Result A	Test Result ¬A	
Condition A	54	10	64
Condition ¬A	47	127	174
	101	137	238

1. What is the:

   a. False negative (Type I error) rate.

   b. False positive (Type II error) rate.

   c. Sensitivity.

   d. Specificity.

   e. Positive predictive value.

   f. Negative predictive value.

2. Is this a valid methodology?

O. How many "true values" are consistent with the result of a typical measurement?

P. What type of considerations would go into the validation of a new method?

APPENDIX A

MISCELLANEOUS RESOURCES

I. ACRONYMS

A.	AAFS	American Academy of Forensic Sciences
B.	ASCLD	American Society of Crime Lab Directors
C.	ASTM	American Society for Testing and Materials
D.	BIPM	International Bureau of Weights and Measures
E.	FQS-I	Forensic Quality Services - International
F.	IEC	International Electrotechnical Commission
G.	IEEE	Institute of Electrical and Electronics Engineers
H.	ILAC	International Laboratory Accreditation Cooperation
I.	ISO	Greek for equal. Not an acronym.
J.	IUPAC	International Union of Pure and Applied Chemistry
K.	JCGM	Joint Committee for Guides in Metrology
L.	NAS	National Academy of Sciences
M.	NIST	National Institute of Standards and Technology
N.	NFSTC	National Forensic Science Technology Center
O.	OIML	International Organization of Legal Metrology
P.	SOFT	Society of Forensic Toxicologists

II. STANDARDS

A. *General requirements for the competence of testing and calibration laboratories* – ISO 17025 (2005) http://www.iso.org/iso/catalogue_detail?csnumber=39883 (purchase required).

B. *National Voluntary Laboratory Accreditation Program Procedures and General Requirements* – NIST HB150 (2006) http://ts.nist.gov/Standards/Accreditation/upload/nist-handbook-150.pdf .

C. *International Vocabulary of Metrology — Basic and General Concepts and Associated Terms (VIM)* – VIM JCGM 200 (2008) http://www.bipm.org/utils/common/documents/jcgm/JCGM_200_2008.pdf .

D. *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 .

E. *The International System of Units (SI)* – JCGM (8th ed. 2008) http://www.bipm.org/utils/common/pdf/si_brochure_8.pdf.

F. *Guidelines for Evaluating and Expressing the Uncertainty of NIST Measurement Results* – NIST TN 1297 (1994) http://physics.nist.gov/Pubs/guidelines/TN1297/tn1297s.pdf .

G. *Guideline for Forensic Science Laboratories* – ILAC G19 (2002) http://www.renar.ro/acreditare_rom/Doc%20ILAC/Ilac-g19ForensicScienceLaboratories.pdf.

H. *Metrological Traceability of Measurement Results in Chemistry, Provisional Recommendation* – IUPAC (2008) http://old.iupac.org/reports/provisional/abstract07/fajgelji_draft_2007-09-18.pdf .

I. *Quality Assurance Standards for DNA DataBasing Laboratories* – FBI (effective July 1, 2009) http://www.cstl.nist.gov/biotech/strbase/QAS/Final-FBI-Director-Databasing-Standards.pdf.

J. *Quality Assurance Standards for Forensic DNA Testing Laboratories* – FBI (effective July 1, 2009) http://www.cstl.nist.gov/strbase/QAS/Final-FBI-Director-Forensic-Standards.pdf .

K. Strengthening Forensic Science in the United States: A Path Forward – NAS (2009) http://www.nap.edu/catalog.php?record_id=12589 .

III. Web sites of metrological and accreditation bodies

A. ISO: http://www.iso.org/iso/home.htm

B. NIST: http://www.nist.gov/index.html

C. ASTM: http://www.astm.org/

D. BIPM: http://www.bipm.org/en/home/

E. OIML: http://www.oiml.org/

F. EURACHEM: http://www.eurachem.org/

G. ASCLD/LAB: http://www.ascld-lab.org/

H. ILAC: http://www.ilac.org/

IV. State weights and measures information, statutes and sites:

A. State weights and measures statutes: http://ts.nist.gov/WeightsAndMeasures/WMLAW.cfm.

B. List State weights and measures programs with contacts and whether NIST certified/accredited: http://ts.nist.gov/WeightsAndMeasures/statelabcontact.cfm

C. Official site for Georgia: http://agr.georgia.gov/00/article/0,2086,%2038902732_0_40992848,00.html

D. Official site for Idaho: http://www.agri.idaho.gov/Categories/WeightsMeasures/metrology.php

E. Official site for New Mexico: http://www.nmda.nmsu.edu/weights-and-measures

F. Official site for North Carolina: http://www.ncagr.gov/standard/Labs/

G. Official site for North Dakota: http://pc6.psc.state.nd.us/jurisdiction/weights.html

V. Useful web based tools

A. NIST Reference on Constants, Units and Uncertainty: http://physics.nist.gov/cuu/index.html

B. NIST Traceability: http://ts.nist.gov/traceability/

C. NIST Engineering Statistics Handbook: http://www.itl.nist.gov/div898/handbook/

D. Elementary Concepts in Statistics: http://www.statsoft.com/textbook/stathome.html?stbasic.html&1

E. Statistics and Science: Monograph Series: http://projecteuclid.org/DPubS?service=UI&version=1.0&verb=Display&handle=euclid.lnms/1215091126

F. Web Pages that Perform Statistical Calculations: http://statpages.org/index.html

G. Short Tandem Repeat DNA Internet DataBase: http://www.cstl.nist.gov/div831/strbase/

H. DNA Advisory Board Quality Assurance Standards for Forensic DNA Testing Laboratories: http://www.cstl.nist.gov/strbase/dabqas.htm

I. EURACHEM Guides and Documents: http://www.eurachem.org/guidesanddocuments.htm

J. Forensic Science Resources on the Internet: http://www.istl.org/03-spring/internet.html

K. DNA Forensic Mathematics: http://dna-view.com/

L. Forensic Statistics and Legal Reasoning: http://www.josephbell.org/

VI. Journals – Free Access

A. Journal of Research of NIST: http://nvl.nist.gov/nvl3.cfm?doc_id=89&s_id=117

B. Measurement Science Review: http://www.measurement.sk/

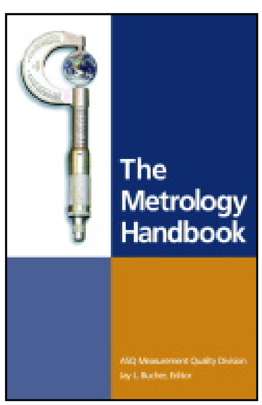
C. Metrology and Measurement Systems: http://www.metrology.pg.gda.pl/

D. Pure and Applied Chemistry: http://www.iupac.org/publications/pac/index/

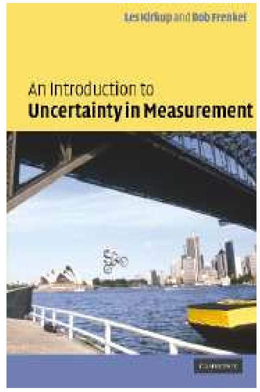
E. The Journal of Philosophy, Science and the Law: http://www6.miami.edu/ethics/jpsl/index.html

VII. Books

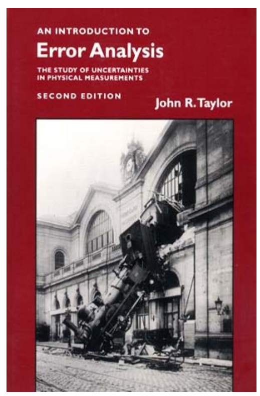
A. *The Metrology Handbook* (ASQ 2004).



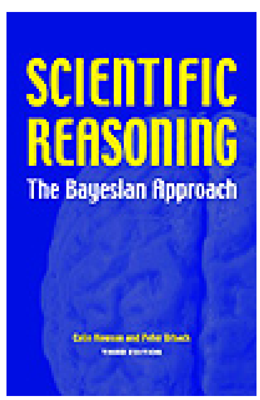
B. Kirkup, *An Introduction to Uncertainty in Measurement* (Cambridge University Press 2006).



C. Taylor, *An Introduction to Error Analysis: The Study of Uncertainties in Physical Measurements* (2^{nd} Ed. 1997).



D. Howson, *Scientific Reasoning The Bayesian Approach* (Open Court 2006).



Coming Soon

Vosk, Emery, Fitzgerald, *Forensic Metrology: A Primer on Scientific Measurement for Lawyers, Judges and Forensic Scientists* (CRC Press – In Preparation)

APPENDIX B

State v Ahmach, King County District Court (Jan. 2008)

IN THE DISTRICT COURT OF KING COUNTY FOR THE STATE OF WASHINGTON

EAST DIVISION, REDMOND COURTHOUSE

STATE OF WASHINGTON, Plaintiff, vs. AHMACH, SANAFIM, ET AL. Defendants	Case No. C00627921, ET AL. ORDER GRANTING DEFENDANTS' MOTION TO SUPPRESS

Each of the Defendants joined in this motion ask that this three judge panel of the King County District Court suppress the Defendants' breath test readings, arguing that the Washington State Toxicology Laboratory (WSTL) engaged in practices which were both fraudulent and scientifically unacceptable. The State, while agreeing that many of the activities of the WSTL were unacceptable, argues that suppression is not the appropriate remedy, both because none of the Defendants' tests were directly affected at any critical point and because the issues raised by the Defendants could be raised before each trier of fact and given their appropriate weight.

For the reasons stated in this Order, the breath tests in each of the Defendants' cases are suppressed.

Findings of Fact

Each of the Defendants herein were arrested for an alcohol related traffic offense, and each submitted to a test of his or her breath at the request of the arresting officer. These tests were performed on the Datamaster or Datamaster CDM machines located throughout King County and Washington.

These instruments operate under the principal of comparing the unknown (the breath of the arrestee) to a known standard of alcohol to measure the amount of alcohol in the breath. There are multiple checks performed by the instrument to ascertain the accuracy of the result. One of the checks is the external standard, which measures the headspace alcohol vapor content of an external simulator solution (field solution). This solution is a mixture of ethanol and water in a known quantity prepared by the WSTL.

These instruments are periodically checked, calibrated and maintained by the Washington State Patrol Breath Test Section (breath test section). For this purpose they also use solutions of ethanol and water prepared to known standards by the WSTL (QAP solutions).

The procedure for preparation of QAP and field simulator solutions is set forth in protocols created and/or promulgated by the State Toxicologist, Dr. Barry Logan. An analyst mixes the solutions according to the protocol, and then each of 16 analysts test the solutions by preparing vials of the mixture and submitting them to headspace gas chromatography along with control vials and blank vials. The results are recorded for each analyst, and ultimately published to the web for access by the public. The analysts then “certify” that they have performed the tests, and that the results as published are correct. These certifications are intended to be used in court in lieu of live testimony by the toxicologists.

This three judge panel has found many irregularities in the preparation, use and documentation of these solutions and tests, as set forth below:

False Certifications

1. Ann Marie Gordon (AMG) became lab manager at WSTL by appointment of Dr. Logan.
2. AMG informed Dr. Logan that her predecessor as lab manager had engaged in a practice of having other toxicologists prepare and test simulator solutions for him and yet certify that he had prepared and tested the simulator solutions.
3. AMG told Dr. Logan that she did not approve of this procedure and was then also informed by Dr. Logan that it was not acceptable for a toxicologist to engage in this practice.
4. Nonetheless, AMG did engage in this practice beginning in 2003. Ed Formoso was a lab supervisor; he prepared and tested simulator solutions for AMG from 2003 to 2007. This involved 56 simulator solution tests.
5. Each test was accompanied by a CrRLJ 6.13 certification that AMG had performed the test and that the test was accurate and correct.
6. Melissa Pemberton was the quality control manager at the WSTL during a part of this time, and knew that AMG was not performing tests but was certifying them.
7. This deception was uncovered after two anonymous tips received by the Chief of the Washington State Patrol.
8. The first was received on March 15, 2007. Dr. Logan was directed by Assistant Chief Beckley to investigate this complaint.
9. Dr. Logan directed AMG and Formoso to investigate the complaint.
10. AMG and Formoso discussed the procedure and agreed that Formoso would no longer perform tests on behalf of AMG.

11. AMG informed Dr. Logan that she did not perform the tests of the solutions but that she signed the forms indicating that she did.

12. AMG and Formoso prepared a report stating that there was no problem with the certifications and that no solution had left the lab with an incorrect solution in 20 years.

13. Dr. Logan, AMG and Formoso knew, or should have known, that this report was incorrect and misleading, but took no steps to correct it or provide for another investigation.

14. Melissa Pemberton had run vials prepared for AMG by Formoso through the gas chromatograph along with her own samples, knowing that these were to be attributed to AMG, and that AMG would sign certificates alleging that she did the tests.

15. Dr. Logan was aware of this, by August of 2007.

16. DR. Logan and Pemberton both testified under oath that no one other than Formoso ever ran tests for AMG.

Defective and Erroneous Certification Procedures

17. The software used to perform calculations for simulator solution worksheets was defective from its inception in that it omitted the fourth data entry from the fourth toxicologist who performed the tests.

18. Beginning in August 2005 a change in the software resulted in a failure to include data from 4 of the 16 toxicologists performing tests in calculations to establish accuracy.

19. Lab protocols require the inclusion of all analysts' data in these calculations.

20. No one checked the software program to ascertain accuracy and compliance with protocols. There was no procedure or protocol propounded to check or verify software used by the WSTL.

21. Analysts were not trained or directed to check the calculations performed by the software.

22. Analysts regularly signed declarations which stated the mean concentration of alcohol in the solutions. These declarations were prepared by support staff, and were not checked for accuracy by the analysts before signing. In at least six instances these declarations were in error. At lest one analyst signed them a second time still reflecting the errors.

Software Failure, Human Error, Equipment Malfunction and Violation of Protocols

23. The software used for calculations to determine the acceptability of simulator solutions was developed by computer programmer(s) within the Washington State patrol and was not subject to rigorous testing and/or checking such that substantial errors resulted and significant data was deleted from calculations.

24. No procedure or protocol within the WSTL required this software to be validated for accuracy or fitness for purpose, and no Lab personnel conducted such testing at anytime, nor verified that the data produced was correct.

25. Errors based on software miscalculations existed within almost all field simulator solution certifications issued between August 2005 and August 2007. At least one QAP solution was similarly affected.

26. When analysts conducted gas chromatograph tests, the machine printed results automatically. These were maintained in the test files. Thereafter (sometimes weeks

after), worksheets were prepared by support personnel detailing the testing results for each toxicologist. Thereafter analysts signed the worksheets to acknowledge their correctness. These worksheets were not checked against the original chromatographs to determine if they were accurate before signing, and incorrect data was in fact inserted into some worksheets. These worksheets were posted to the web and relied upon in determining the accuracy and precision of the breath testing machines in the field.

27. Declarations by toxicologists for certification of the solutions are prepared by support personnel and then given to analysts to sign, sometimes weeks after the actual testing. These were not checked against chromatographs or worksheets to insure accuracy. There were at least 150 instances of similar non-software related errors committed by analysts and revealed in the record. These include:

a. Entering incorrect data into certification spreadsheets for use in calculations to determine mean solution values and compliance with protocols.

b. Entering incorrect test values for controls.

c. Entering data for the wrong solutions into certification spreadsheets.

d. Signing declarations indicating testing of the solution prior to the solution even being prepared.

e. Signing declarations indicating that a solution had been tested before the testing had taken place.

f. Incorrect dates for testing and/or signing of declarations.

28. The WSTL was equipped with several gas chromatograph machines for use by the analysts. A machine that malfunctioned was not repaired or maintained adequately

and this resulted in different operational and measurement characteristics and abnormal variations in readings. The machine remained on line for some time even though individual toxicologists knew that it was not functioning properly. Once repaired this abnormality disappeared.

Improper Evidentiary Procedures

29. In 2004 the Washington State Patrol conducted an internal audit of the WSTL. The report included the following conclusions:

a. The WSTL was noncompliant with policies and procedures in 8 major categories.

b. The simulator solution logbooks were not properly kept.

c. The required self audits were not performed.

d. AMG indicated that she did not have time to follow WSP policies and would not do so.

e. "WSP policies and required procedures appear to be of secondary concern to Lab personnel....Accurate recordkeeping and quarterly auditing as required by patrol Policies and CALEA standards is severely deficient."

30. In 2007 another internal audit was conducted by the Washington State Patrol. The report included the following conclusions:

a. "The department is unnecessarily exposed to litigation due to insufficient documentation and disregard for evidence handling policies and procedures."

b. "Mandatory audits are not being completed.... Non-standard evidence handling procedures and insufficient documentation to ensure the same...and failure to perform required audits jeopardizes operational performance as well as CALEA accreditation. "

Inadequate and Erroneous Protocols and Training

31. The accuracy of breath alcohol measurements is determined by the use of simulator solutions. These must be accurately prepared and certified as such to gain the trust and confidence of the courts and public.

32. Accuracy of these solutions is assured by the adherence to proper protocols for their preparation and use.

33. Contrary to protocol requirements, toxicologists were trained to discard data generated by the tests if any single data entry lay outside the range for the mean value of the solution as dictated by the protocol. This tended to create a testing system that would not fail a solution as every value outside the range was discarded and only those that were within the accepted range were included in the calculations of accuracy.

34. Discarding of data is appropriate in some circumstances where identifiable reasons exist or where there is appropriate statistical justification (outliers). However, a decision to discard data must be governed by appropriate protocols and must be properly documented so that these decisions can be reviewed. Such a protocol was not promulgated until this legal proceeding was well underway, and documentation was not required or provided.

35. Several toxicologists discarded data without identifiable or statistical reasons for doing so. Inadequate or no documentation was provided, so that in those situations this Court cannot determine why data was discarded.

36. At least one toxicologist was not taught that testing of simulator solutions followed different procedures than testing of other materials, and conducted multiple tests, discarding the results of at least one test.

37. Protocols for solution preparation and machine testing were contradictory or inconsistent, resulting in field solutions being used for QAP testing in some cases.

Impact on Tests Conducted In the Field

38. Field solution #2018 was never properly certified due to errors committed by the analyst. This solution was used as the external standard in 2,018 tests.

39. Field solution #2019 was never properly certified due to similar errors committed by the same analyst. These two batch errors were likely caused when the analyst switched data. This solution was used as the basis for QAP's performed on at least 39 breath test machines. There were approximately 7,928 tests conducted on the affected machines.

40. QAP batch solution #06028 was certified after data was discarded improperly. QAP procedures were performed on 32 Datamaster machines using this solution. This had an impact on 3,445 tests.

41. Field solution #05008 was used as a QAP solution to test and calibrate the Datamaster. Though, perhaps, not a violation of protocol since the protocols were in conflict, Dr. Logan conceded that field solutions were never intended to be used for the QAP process. This solution was improperly certified by AMG. If the data from her tests were removed, the solution has a mean alcohol concentration of .1022, outside the acceptable range for QAP solutions. The tests conducted using machines tested and calibrated with this solution number 1,679.

42. Field solution batch #06003 was used as a QAP solution. This solution had a mean alcohol concentration of .1024, outside the range deemed acceptable for QAP solutions. Two machines were tested using this solution, affecting 392 individual tests.

43. Field solution #06048 was qualified using software which provided incorrect results. When correct figures are computed, it was determined that the solution would not have qualified as a QAP solution. At least one Datamaster QAP was performed with this solution, affecting 21 individual tests.

44. This same solution was also used as a field solution, but when proper calculations are made, it is apparent that it would have affected all tests conducted using this machine. However, the number of tests affected has not been determined.

45. QAP solution #06037 was certified using software that incorrectly calculated the equivalent vapor concentration. The machines calibrated using this solution affected 2,691 individual breath tests.

46. Field solution #06043 was tested by one analyst using a defective gas chromatograph. The test should have been repeated to determine accuracy. The number of individual test impacted by this has not been ascertained.

47. Not all (or possibly any) of the defective solutions noted above would have resulted in substantial changes in every test result. Some test results would be of greater importance than others if they are at or near the absolute standards for violations created by statues, ie. .02, .04, .08, and .15. However, every test conducted with an improperly certified or defective solution is affected in some way.

Nondisclosure of Machine Bias

48. All measuring machines have some bias, and Datamaster breath test machines have bias which is identified in the QAP process.

49. This bias is not determinable without testing; sometimes creating readings lower than actual and sometimes higher.

50. The bias of any particular machine can be determined from the information created during the QAP process by applying mathematical formulas and calculations. This information is not readily available to the public, though it is published on the web. Due to the complexity of the calculations and formula involved, few in the legal community are aware of this bias. The Breath Test Section of the Washington State Patrol does, however, provide this information to attorneys and defendants when requested.

51. The machine bias information could be easily made available to the defendants, attorneys and public by the State Toxicologist.

Analysis

BAC Admissibility Post <u>Jensen</u>

The Washington legislature conveyed its "frustration with the inadequacy of previous attempts to curtail the incidence of (Driving Under the Influence) DUI" with the adoption of SHB 3055[1] in 2004. <u>City of Fircrest v. Jensen, 158 Wn.2d 384, 388 (2006)</u>. Central to SHB

---

[1] In part, the legislature indicated its intent in the adoption of SHB 3055 as follows:
"The legislature finds that previous attempts to curtail the incidence of driving while intoxicated have been inadequate. The legislature further finds that property loss, injury, and death caused by drinking drivers continue at unacceptable levels. This act is intended to convey the seriousness with which the legislature views this problem. To that end the legislature seeks to ensure swift and certain consequences for those who drink and drive.

To accomplish this goal, the legislature adopts standards governing the admissibility of tests of a person's blood or breath. These standards will provide a degree of uniformity that is currently lacking, and will reduce the delays

3055 were amendments to RCW 46.61.506, by which the legislature sought to curtail pretrial motions seeking the suppression of breath tests in DUI cases. As amended, RCW 46.61.506 required that trial courts assume the 'truth of the prosecution's... evidence and all reasonable inferences from it in a light most favorable to the prosecution." RCW 46.61.506(4)(b). While the amendments would still allow defendants to challenge the reliability or accuracy of breath tests, those challenges would "not preclude the admissibility of the test once the prosecution ... has made a prima facie showing" of each of eight basic admissibility requirements set forth in the statute. RCW 46.61.506(4)(a). Ultimately then, SHB 3055 constituted a legislative attempt to eliminate the trial court's role as the gatekeeper[2] for a critical piece of evidence in DUI prosecutions.

Thus, when the Washington Supreme Court considered this issue in Jensen, supra, the court could have found that the legislation violated the inherent right of the judicial branch to control its own court procedures, i.e., a violation of the Separation of Powers doctrine. Instead, the Court determined that it could harmonize RCW 46.61.506, as amended, with the rules of evidence and give effect to both. Jensen, 158 Wn.2d at 399. The court held that, once the prosecution had met its prima facie burden under RCW 46.61.506(4), the breath test thereafter became "admissible," meaning that the court could still serve in its role as the gatekeeper under the applicable rules of evidence. Id. By analogy, the Jensen court referenced DNA testing:

---

caused by challenges to various breath test instrument components and maintenance procedures. Such challenges, while allowed, will no longer go to admissibility of test results. Instead, such challenges are to be considered by the finder of fact in deciding what weight to place upon an admitted blood or breath test result."
Laws of 2004, ch. 68.

[2] A trial court is said to be the "gatekeeper" for the admissibility of evidence under both the Frye test (Frye v. United States, 293 F. 1013 (D.C. Cir. 1923)) and under the standard articulated in Daubert v. Merrell Dow Pharmaceuticals, Inc., 509 U.S. 579 (1993); State v. Copeland, 130 Wn.2d 244, 259-260 (1996). "In Daubert, the Supreme Court held that a trial judge should act as a "gatekeeper" to ensure that all scientific evidence admitted is both relevant and reliable." Reese v. Stroh, 74 Wn. App. 550, 559 (1994). The court also acts as the gatekeeper when it rules on motions to suppress scientific evidence under ER 403 or ER 702.

> In the DNA analogy, DNA admissibility has been accepted under Frye[3]; however, challenges to the weight of the DNA evidence, including laboratory error; the size, quality, and randomness of Federal Bureau of Investigation (FBI) databases, and the methodology and practices of the FBI in declaring a DNA match, are subject to ER 702 admissibility as determined by the trial court.

Jensen, 158 Wn.2d at 397. Continuing this analogy to the cases herein, the trial court's determination that the prosecution had, prima facie, met the requirements of RCW 46.61.506(4), would be comparable to acceptance under Frye, meaning that the court would then move on to consideration of any rules of evidence that might be applicable.

ER 702 and Laboratory Evidence

A breath test reading is not admissible absent expert testimony, either in person or by affidavit as allowed by CrRLJ 6.13(c)[4]. Pursuant to ER 702, however, an expert may only testify "if scientific, technical, or other specialized knowledge will assist the trier of fact to understand the evidence or to determine a fact in issue." In a criminal prosecution, a post Frye analysis of the admissibility of expert testimony under ER 702 is a consequential activity with independent force and effect. "In this state ER 702 has a significant role to play in admissibility of scientific evidence aside from Frye." State v. Copeland, 130 Wn.2d 244, 259-260 (1996).

[3] Frye requires that the court determine whether (1) the scientific theory has general acceptance in the scientific community, (2) the techniques and experiments that currently exist can produce reliable results and are generally accepted by the scientific community, and (3) the laboratory performed the accepted scientific techniques in the particular case. Frye v. United States, Supra.

[4] A breath test technician must testify that the BAC Verifier Datamaster or Datamaster CDM was tested, certified and working properly on the date of the test, and a state toxicologist must testify that the simulator solution was properly prepared and tested. Both would also have to testify that each activity was performed in conformance with the rules established by the Washington State Toxicologist. RCW 46.61.506(3); CrRLJ 6.13(c).
The Defendants here have sought suppression of their breath tests based upon the failure of the WSTL to properly prepare, test and certify simulator solutions. The Defendants have not raised any issues relating to the Washington State Patrol Breath Test Section or Breath Test Technicians.

Under Jensen, therefore, after the prosecution has met its prima facie burden for the admission of a BAC reading, a trial court must engage in a meaningful review of the admissibility of the BAC evidence involving, under ER 702, a two part test. State v. Cauthron, 120 Wn.2d 879, 890 (1993). As in Copland, supra, the Cauthron court was concerned with the admissibility of DNA evidence:

> The 2-part test to be applied under ER 702 is whether: (1) the witness qualifies as an expert and (2) the expert testimony would be helpful to the trier of fact. Part 2 of this standard should be applied by the trial court to determine if the particularities of the DNA typing in a given case warrant closer scrutiny. If there is a precise problem identified by the defense which would render the test unreliable, then the testimony might not meet the requirements of ER 702 because it would not be helpful to the trier of fact.

Cauthron, 120 Wn.2d at 890. In each of the following cases, the Supreme Court engaged in both a Frye analysis and an ER 702 review of challenged forensic laboratory conclusions. In each case discussed, the court began with the proposition that the "determination of whether expert testimony is admissible is within the discretion of the trial court. Unless there has been an abuse of discretion, this court will not disturb the trial court's decision." Cauthron, 120 Wn.2d at 890. In each case the trial court admitted the scientific evidence and none of the ER 702 challenges to the trial court decisions were overruled, both for the factual reasons noted for each below, and because in each case the court was upholding a discretionary ruling of the trial court.

- In State v. Cauthron, supra, the court noted that the defense had only presented "potential problems" with the DNA evidence. Moreover, the court noted that "the defense presented its own experts to rebut the State's conclusions. Dr. Ford and

Dr. Libby both testified that they found the autorads in this case inconclusive, and discussed their reasons at length. In addition, they each pointed out the possible pitfalls of DNA testing, such as degradation, starring, cross contamination, etc., and the lack of controls employed in the testing procedure. The jury was presented with a balanced picture of the DNA evidence[5]." Cauthron, 120 Wn.2d at 899.

- In State v. Kalakosky, 121 Wn.2d 525 (1993), the court quickly dealt with the two errors cited by the defense. (1) "The defense asserts that semen samples taken from the C.F. crime scene were spilled in 'close working proximity to samples of defendant's blood'. The record does not support this". Kalakosky, 121 Wn.2d at 540. (2) "The defense also alleges that there was evidence of a mislabeled autoradiograph which compromised the reliability of the DNA testing. This also is unsupported by the record." Id.

- In Copeland, supra, the court considered the admissibility of lab results which had been challenged for a lack of external testing of lab procedures and for allegedly simplistic proficiency testing procedures. In dismissing these challenges, the court noted that "while a completely independent audit may be ideal, there was no evidence that the FBI procedures compromised the test results in this case." Copeland, 130 Wn.2d at 271. The court concluded that the "issues of laboratory error and lack of proficiency testing can be and were the subject of cross-

---

[5] The Cauthron court ultimately reversed the trial court, not for lab error, but because a critical underlying assumption for the admissibility of DNA testing was absent. "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." Cauthron, 120 Wn.2d at 907.

examination and defense expert testimony at Copeland's trial. Id.; See also, State v. Cannon, 130 Wn.2d 313 (1996).

Thus, in each of the above cases dealing with potential lab errors and poor lab procedures, the errors and poor procedures were relatively insignificant. Moreover, the Supreme Court stressed the importance of a trial court's role in evaluating lab evidence under the mandates of ER 702.

In Kalakosky, while the court noted that alleged infirmities in the performance of a test will usually to go to the weight of the evidence, not its admissibility, it also stated that:

> If the testimony before the trial court shows that a given testing procedure was so flawed as to be unreliable then the results might be excluded because they are not "helpful to the trier of fact". The issue of human error in the forensic laboratory is analyzed under ER 702 and is not a part of the Frye test....

Kalakosky, 121 Wn.2d at 541. See also, Cannon, 130 Wn.2d at 325; and Copeland, 130 Wn.2d at 270. That this is still the standard in DUI cases post Jensen is reflected in Justice Madsen's concurrence in City of Seattle v. Ludvigsen, 2007 Wash. LEXIS 953 (2007):

> When deviations from additional testing procedures or machine maintenance protocols are so serious as to render test results unreliable, a court has discretion to exclude them in accordance with the rules of evidence.

Ludvigsen, at page 35.

The State argues a violation of protocols by the WSTL could not provide any basis for suppression of breath tests, citing State v. Mee Hui Kim, 134 Wn. App. 27 (2006). Kim, however, does not stand for the proposition that a breath or blood test may never be suppressed for a violation of WSTL protocols under ER 702. The defendant in Kim did not contend that the WSTL failed to comply with a protocol; rather the defendant in Kim argued that the State had failed to *show* compliance with a protocol:

> Specifically, Kim points to the State's failure to show that preparation of the volatile standards in the "Alcohol Standard Logbook" met the requirements in the Head Space GC Protocol.

Kim, 134 Wn. App. at 35-36. Ann Marie Gordon, testifying at the Kim motion hearing, stated that the protocol had been complied with and that the logbook was available at the lab for defense review. Upon these facts the trail court held that the State had shown compliance with the WAC and that the defense could (when, after the motion hearing they had been able to review the logbook) renew their motion to suppress. Kim, 134 Wn. App. at 36-37. Thus, trial courts are still able to weigh the failure of the WSTL to follow its own protocols in a motion to suppress under ER 702.

In each of the Defendants' cases herein, the defense cannot point to specific errors directly compromising the breath test results at critical BAC levels. For this reason the State argues that this court should decline to suppress the results of the breath tests and should instead admit the evidence at trial and allow the trier of fact to weigh each of the issues raised. While the State's position is generally preferable when disputes arise relating to the quality of scientific evidence, it is not always the last word on the subject. Indeed, if the court were always to admit questionable evidence at trial, ER 702 would serve little purpose. Here we find, for the reasons documented in this court's findings of fact and more fully explained below, that the decision to suppress or admit tips considerably in favor of suppression.

Under the current statutory scheme, a charge of DUI is most commonly proven by two different means; proving that an individual drove a motor vehicle while under the influence of or affected by intoxicating liquor, or by proof that the person had, within two hours after driving, an alcohol concentration of 0.08 or higher as shown by analysis of the person's breath[6]. RCW 46.61.502 (1). Proof of DUI via analysis of the persons breath is considered a per se violation, i.e., the state is not required to show that the defendant was affected by the alcohol, merely that the level of alcohol in the defendants breath was at or above 0.08. Thus, a crime which carries a potential sentence of one year in jail; carries a mandatory minimum of some amount of jail time, and which will result in the mandatory loss of the privilege to drive a motor vehicle, may be proved by evidence from an instrument alone.

The 0.08 BAC level is not the only critical level for breath alcohol which has been set by the legislature. The first critical level is 0.02, the level at which a person under the age of 21 may be convicted of Driving or Being in Physical Control of a Motor Vehicle After Consuming Alcohol. RCW 46.61.503. The next critical breath alcohol level is 0.04, the level at which a commercial driver will lose his or her commercial drivers license (CDL) for one year. RCW 46.25.090; RCW 46.25.120. Finally, in a DUI prosecution, in addition to the 0.08 breath alcohol level, the 0.15 level is also critical. A breath alcohol level of 0.15 or above carries greater mandatory minimum sentencing requirements. RCW 46.61.5055. Moreover, for breath tests

[6] The state may also prove the charge of DUI by proof that the defendant was under the combined influence of liquor and any drug or by proof that the defendant's blood alcohol concentration was 0.08 or higher. RCW 46.61.502 (1).

registering above 0.02, 0.04 and 0.08, an individual may lose his or her privilege to drive without the benefit of a prior hearing[7]. RCW 46.20.3101; RCW 46.25.120.

Thus, even errors in the range of 1 or 2% can have a profound effect on a breath test reading. Nonetheless, each expert witness who offered testimony[8] stated that there was not a process or a machine that would not insert some amount of inherent error in any result. That is also the case with the Datamaster and Datamaster CDM. In the process of breath test instrument calibration, the protocols indicate that breath test instrument is still functioning properly if it is accurate to within +/- 5%, and if the precision of the readings stand at +/- 3%[9]. Rod Gullberg testified that the lack of accuracy in a breath test machine is referred to as "bias." A breath test machine normally has a bias of 1-2%, with the smaller fraction of the machines registering a bias of 5% or less[10]. The breath test program is not, however, set up to account for any of the potential bias inherent in a breath test machine[11]. Thus, a process that already allows potential bias in each reading only underscores the importance of ensuring that the WSTL eliminates all other possible sources of error.

Throughout Washington State, over 40,000 breath tests are administered annually. In light of the importance of each one of these tests for the state and for individual defendants, it is vital that each aspect of the breath test program operate effectively. As stated in the findings, the WSTL prepares and tests both field simulator solutions and quality assurance procedure

---

[7] In the case of a 0.04 reading, a CDL is lost. In each situation the defendant may request a hearing prior to revocation.
[8] The court heard testimony from the following expert witnesses: Rod Gullberg, Dr. Barry Logan, Dr. Ashley Emery and Dr. Nayak Pollisar.
[9] The WAC defines accuracy and precision as follows: "accuracy" means the proximity of a measured value to a reference value; "precision" means the ability of a technique to perform a measurement in a reproducible manner. WAC 449-16-030 (1) & (10).
[10] The bias allowed in the protocols, however, does not include improper procedures or mistakes.
[11] For instance, readings are not adjusted at any of the critical levels to account for actual of for potential bias, nor are defendants informed of the potential bias before or during trial.

simulator solutions. These solutions serve as a critical check on breath test instruments to ensure that each will provide accurate and precise breath alcohol readings. The CrRLJ 6.13 certificates, or a toxicologist's in-court testimony, allow a breath test technician to "close the loop" and testify that the breath test reading was correct.

A Culture of Compromise

The Cauthron, Kalakosky and Copeland cases, discussed above, generally dealt with questions of lab mistakes and process errors. While many of our findings concern lab mistakes and process errors, the remaining findings indicate that the problems in the WSTL are much more pervasive.

Generally, our concerns regarding the WSTL fall into three general categories:

1. The failure to pursue the ethical standard which should reasonably be expected of an agency that operates as an integral part of the criminal justice system;
2. The failure to establish procedures to catch and correct human, and software and machine errors within the lab; and
3. The failure to pursue the rigorous scientific standards which should be reasonably expected of an agency that contributes a key component of critical evidence that may, almost standing alone, result in a criminal conviction.

Ethical Compromises

Ann Marie Gordon falsely signed CrRLJ 6.13 certifications under penalty of perjury indicating that she prepared and tested field simulator solutions and that the solutions were found to conform to the standards established by the State Toxicologist. This and other ethical compromises documented in the findings adopted in this order may at the same time be viewed as both petty and alarming. The ethical compromises were petty because they were frustratingly unnecessary, and alarming because the WSTL exists primarily to provide accurate information to state trial courts[12]. It is, therefore, reasonable to expect that those employed in an office with such a direct link to courts, whose primary duty is the discovery of the truth, would fully understand the importance of truth in all of their activities. The State has argued that there isn't any evidence that Ann Marie Gordon ever actually testified in court that she had prepared and tested a simulator solution. Yet, CrRLJ 6.13 exists to allow the admission of simulator solutions (via affidavits) in the absence of direct court testimony by the toxicologist who prepared the solution. We do not know whether any false Ann Marie Gordon CrRLJ 6.13 certificates were ever used in court in lieu of live testimony, but considering the number of DUI trials, it is more than likely that some were.

There are several other factors that highlight the disturbing nature of this practice. This was a procedure which:

- Ann Marie Gordon herself had specifically recognized was inappropriate;
- violated the protocols of the WSTL;
- required that she not only state that she performed an activity which she did not perform but also that she sign an affidavit to that effect under penalty of perjury;

[12] The WSTL was created to provide forensic information to prosecuting attorneys as well as coroners and medical examiners. Prosecuting attorneys will, of course, request information from the WSTL in the hope that it will assist in the prosecution of anyone who may be guilty of committing a crime. In the case of breath alcohol testing, the link to trial courts is strong because the WSTL runs essentially independent of specific requests from individual prosecuting attorneys.
The WSTL was specifically established by RCW 68.50.107:

> "There shall be established in conjunction with the chief of the Washington state patrol and under the authority of the state forensic investigations council a state toxicological laboratory under the direction of the state toxicologist whose duty it will be to perform all necessary toxicologic procedures requested by all coroners, medical examiners, and prosecuting attorneys."

- required the active participation of at least one other member of the WSTL (Edward Formoso) in the fraud (but we have also found that this pernicious fraud ultimately required the participation of toxicologist Melissa Pemberton and perhaps others)[13]; and
- set the ethical tone for the entire toxicology lab[14].

While such fraud can never be justified by necessity, it is, nonetheless, baffling to consider the risk the toxicology lab was willing to take for little, if any, gain. If Ann Marie Gordon never testified in court that she prepared and tested a simulator solution, and if this means that she, perhaps, never intended to so testify, why was she so ready to commit perjury by signing false certifications?

The State Toxicologist, Dr. Barry Logan, is ultimately responsible for the WSTL, and he bears a good deal of the responsibility for its shortcomings. He hired and supervised Ann Marie Gordon. Ms. Gordon testified that she continued to "test" solutions and sign the CrRLJ 6.13 certificates because she believed Dr. Logan wanted her to. Dr. Logan testified that he had been told in 2000 by Ms. Gordon that her predecessor in the WSTL had fraudulently signed CrRLJ 6.13 certificates when he was manager of the WSTL. Yet, not only did Dr. Logan fail to detect that this same fraudulent procedure was occurring from 2003 to 2007, but he also professed not to know that toxicologists even signed CrRLJ 6.13 certificates. Because of this ignorance, he testified that he did not understand the meaning of the first tip that came into the State Patrol. The tip indicated that "Simulator solutions are being falsified as far as the

[13] Although we cannot know with certainty whether this fraud was known to the other members of the WSTL, we believe that it is unlikely that anyone working in such a small office could have failed to see that one of their members was failing to test a solution and that, nonetheless, her name would appear on the paperwork they all had to sign indicating that they had each completed their testing.

[14] This conclusion is not meant to indicate that all members of the toxicology lab engaged in unethical practices. It is rather, a comment on the culture of the office itself. If the top of the chain of command engages in questionable practices, it should not surprise anyone to find that this poor behavior has infected the culture of the entire office. Again however, we caution anyone from making any specific conclusions about employees of the WSTL. Good people are quite capable of resisting poor behavior, even if a poor example is set at the top; and during the course of this motion we heard the testimony of many competent, dedicated and ethical people from the WSTL.

certification." Thereafter, in a situation screaming with irony, Dr. Logan assigned the perpetrator of the fraud, Ann Marie Gordon, the task of investigating the tip. To complete the circle, Ms. Gordon enlisted the assistance of lab supervisor Ed Formoso, her co-conspirator in the fraud, as her co-investigator. While they both ended their fraudulent practice at the time the first tip was received, their investigation also concluded that no fraud was occurring.

While it is not clear from the testimony of the various parties, just when Dr. Logan knew of the fraud, he should have known after the first tip. As previously stated, it is most likely that everyone in the WSTL was fully aware of the fraud, and if 16 toxicologists knew, why didn't Dr. Logan? When informed that the certifications were being falsified, why didn't he consider the possibility that his current lab manager was engaging in the same activity that had occurred a few years before? Why was Ann Marie Gordon assigned the task of investigating the tip? While these questions may never be answered, they cast a long shadow over Dr. Logan's ability to serve as the State Toxicologist.

Systemic Inaccuracy, Negligence and Violation of Scientific Principals

Dr. Nayak Polissar, an expert called by the State, testified that only superior methods will ensure accuracy, and that the accuracy and precision necessary for a particular laboratory task is dependent upon the particular use intended for the final product. As stated by the National Institute of Standards and Technology (NIST), "accuracy… is judged with respect to the use to be made of the data." NIST Special Publication 260-100, 2 (1993).

Data Transfer

When each of the 16 toxicologists tested simulator solutions, the data from their tests was recorded on documents known as chromatograms. The data was thereafter transferred to worksheets, a problematic step, unless the WSTL required a review to ensure that the data was correctly transferred. The WSTL did not require that the data transfer be checked, and toxicologists signed certifications which were unverified and later found incorrect. Many errors in diverse areas were subsequently discovered.

Computer Software

The computer software used to enter and calculate simulator solution lab results on the worksheets was not created by an individual with the requisite knowledge and skill necessary to ensure that the data was correctly analyzed and recorded. Moreover, no one checked the software to determine if it was operating properly. Nor was this a mistake that one can charge to an individual employee. The WSTL itself never considered that it was necessary to check the software to ensure that it was fit for its purpose. The software contained errors which were not revealed until the WSTL came under close scrutiny because of the Ann Marie Gordon investigation.

Malfunctioning Gas Chromatograph

The WSTL suffered through a time period during which a gas chromatograph machine was malfunctioning. During this period of time, the gas chromatograph could, under certain circumstances, provide incorrect readings. The WSTL chose to ignore rather than address this issue for a considerable period of time.

Thousands of Tests Affected

Literally thousands of breath tests performed in recent years were affected through a multiplicity of errors in the toxicology lab. A very brief recitation of the errors include: the improper rejection of data; erroneously switched data; the use of field simulator solutions to conduct quality assurance procedures; the use of software that improperly computed data and that improperly ignored the data of the last four of the toxicologists providing data for field simulator solutions; and, the use of simulator solutions that were outside of the allowable range.

Rod Gullberg effectually ran the breath test section for the Washington State Patrol for 25 years. Mr. Gullberg, who, along with Trooper Ken Denton, completed a lengthy review of the solution preparation worksheets from the WSTL, is also well acquainted with the WSTL and its processes. In his opinion, the problems in the WSTL are not the result of bad faith. Instead, Mr. Gullberg believes that the WSTL failures are the result of carelessness and complacency.

Motion to Suppress Granted

While we agree that trail courts should generally admit scientific evidence if it satisfies the requirements of Frye, we also agree that trial courts should thereafter engage in a meaniningful ER 702 analysis, as we have here, when the circumstances require. Having done so, we conclude that, under ER 702, the work product of the WSTL is sufficiently compromised by ethical lapses, systemic inaccuracy, negligence and violations of scientific principals that the WSTL simulator solution work product would not be helpful to the trier of fact[15]. This litany of problems is indicative of a pervasive culture which has been allowed to exist in the WSTL. In this culture, the WSTL compromises the accuracy of the work product. Accuracy becomes secondary to the accomplishment of the work itself. Thus, because of this culture of the expedient, the WSTL has lost its effectiveness.

---

[15] Although many of the problems within the WSTL are of a general nature, our decision today concerns only the simulator solutions prepared and tested by the WSTL. Our decision does not, therefore, directly relate to any of the other work of the WSTL.

This conclusion is especially troubling because of the critical role the WSTL plays in combating the crime of DUI. The criminal justice system is appropriately assigned the task of discovering the truth. Simply stated, without the reliable evidence that a correctly functioning breath test instrument can provide, the discovery of the truth in DUI cases suffers; the innocent may be wrongly convicted, and the guilty may go free.

We wish to emphasize that our decision to suppress today results from the unique multiplicity of WSTL problems highlighted during this motion. Because the identified problems are multiple and diverse, and because the WSTL may find it difficult to prove, in any reasonable manner, that they have corrected each individual problem, we are not able to indicate with specificity, each correction required.

Therefore, while we provide a list of our concerns below, we emphasize that the WSTL is not required to show that each has been corrected. Any one or two problems, standing alone, would not likely have resulted in suppression.

While the WSTL has attempted to modify its practices and procedures as a result of many of the problems noted in the findings herein, and improvements have been made,[16] additional effort is required.

Ethics

The WSTL has not been able to explain how Ann Marie Gordon and Ed Formoso (and perhaps the lab manager prior to Ann Marie Gordon), over a multiple year period, decided that it was acceptable to engage in a practice of falsely signing CrRLJ 6.13 certificates. We are not persuaded that this fraudulent activity should simply be laid at their feet. This apparently long

[16] Indeed, in reaction to a continuing series of discoveries, the State Toxicologist, Dr. Barry Logan amended protocols several times within a recent three month period.

standing ethical lapse is more likely a symptom of a greater problem; a WSTL culture that was tolerant of cut corners.

Errors

While the WSTL has made several policy changes to deal with many of the prolific errors within the WSTL, it has not been able to point to the reasons for what Rod Gullberg stated was a sense of complacency in the WSTL. The WSTL has, to date, simply corrected the systemic errors that have been called to its attention or were discovered as a result of a review of other problems called to its attention. The WSTL must establish procedures that, in the years ahead, ensure that their processes are double checked for accuracy[17].

Forensic Science

The State appropriately relies on the WSTL to produce (as is the case with the simulator solutions) and analyze evidence. The WSTL was not created, however, as an advocate or surrogate for the State. While the WSTL will always assist the State, it must never do so at the cost of scientific accuracy or truth.

In City of Seattle v. Clark-Munoz, 152 Wn.2d 39 (2004), the Supreme Court agreed with the statement that:

> 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.

---

[17] Here we use the word accuracy in its colloquial, non-scientific sense. By the use of the word accuracy, we mean that the WSTL must establish a system which ensures reliability appropriate to the importance of the purpose of each specific task.

Clark-Monoz, 152 Wn.2d at 47. Although the Clark-Monoz holding has been brought into some question as a result of the ruling in Jensen, supra, the proposition that robust scientific standards are expected in the WSTL still remains. And while Rod Gullberg testified that, after the changes made in the WSTL in the fall of 2007, he now has more confidence in the WSTL, more work is required. In the summer of 2008 the WSTL plans to adopt the General Requirements for the Competence of Testing and Calibration Laboratories, ISO/IEC 17025:1999(E), promulgated by the International Organization for Standardization. These standards are neither required for a toxicology laboratory, nor are they a panacea for the past and current problems in the WSTL. Their adoption, however, is likely to move the WSTL a long way toward the type of reliable forensic science which should be expected of a state toxicology lab.

Conclusion

We hold that, under ER 702, the work product of the WSTL has been so compromised by ethical lapses, systemic inaccuracy, negligence and violations of scientific principals that the WSTL simulator solution work product would not be helpful to the trier of fact. The State, perhaps expecting the suppression of some of the work product of the WSTL, has asked this panel to be as specific as possible in our ruling. Specificity is made difficult, however, because of the nature of the problems identified. The State may, therefore, request that this panel reconvene at such time that the State believes it has sufficient evidence that the WSTL has adequately addressed the issues noted in this Order[18].

[18] The alternative, of course, is to seek the admission of breath test evidence before each individual judge who adopts this ruling and then, when the defendants raise the issue, argue case by case that the WSTL simulator solutions currently meet the requirements of ER 702.

Dated this 30th day of January, 2008

Judge David Steiner

Judge Darrell Phillipson

Judge Mark Chow

APPENDIX C

State v Berry, Pierce County District Court (Dec. 2009)

IN THE DISTRICT COURT OF THE STATE OF WASHINGTON
FOR PIERCE COUNTY

STATE OF WASHINGTON,

Plaintiff,

vs.

Leota Berry

Defendant.

No. 9YC000264

OPINION AND ORDER IN RESPONSE TO DEFENDANT'S MOTION TO DISMISS

INTRODUCTION

Defendant brings this motion to suppress the breath analysis offered by the state. Other defendants have subsequently joined this case as co petitioners. The court has reviewed the court files, heard argument, and reviewed the Defendant's Motion to Suppress and the State's Response to Defendant's Motion to suppress and the state's Supplemental Response.

Defendant, Letoya Berry, is charged with Driving Under the Influence of Intoxicants (DUI). The complaint claims that the defendant's breath alcohol was equal to, or exceeded, .08 within two hours of driving, as shown by an analysis of defendant's breath pursuant to R.C.W. 46.61.506 and that she had an alcohol concentration in excess of .015, thereby invoking the enhanced sentence provisions of RCW 46.61 5055 and contrary to RCW 46.61.502 (1)(a),(b),and (c).

The breath readings submitted are .200 and .230 grams of alcohol per liter of breath. Defendant argues the measurements of breath alcohol provided by the Washington State Department of Toxicology Lab (herein after Toxicology Lab) are inadmissible on account of failures to observe scientific protocols in measuring and recording Quality Assurance Procedures and field tests, and a routine of systematic and scientifically unacceptable practices. Defendant also alleged perjury and conspiracy by employees of the toxicology Department.

For purposes of responding to this motion to exclude the court has opted to focus on the scientific aspects that the Court believes are controlling in this case and the companion cases. The Court is electing this approach in recognition of the ongoing reforms adopted by the Toxicology Lab resulting in accreditation for ISO standard 17025 awarded in November 2009 and the personnel changes that have occurred in the past twenty months.

ISSUE:
WHAT STANDARD OF REVIEW IS APPLICABLE TO ADMIT BREATH TEST SCIENTIFIC EVIDENCE OFFERED BY THE STATE TOXICOLOGH LAB?

Over forty years ago in Ward v. J.C. Penny Co., 67 Wn.2d 858,860-61,410 P.2d. 614,616 (1966) the court commented that "the increasing complexity of our society requires a greater use of expert testimony." What was true then is especially true today. As our knowledge base expands at accelerating rates, reliance on experts and machines has become commonplace. Indeed society's growing dependence on technology and experts to build and operate machines to accomplish tasks is impossible to arrest. Because of society's deference to technology, experts, and machines, the rules concerning scientific evidence continue to evolve. However there are some general concepts that are applicable in evaluating the admissibility of scientific evidence

While forensic science is distinct from research science some may believe that a lesser standard is acceptable. Such a conclusion would be erroneous. The authors of National Research Council, *Strengthing Forensic Science in the United States: A Path Forward* maintain quite the contrary at 113:

> In day-to-day forensic science work, the process of formulating and testing hypotheses is replaced with the careful preparation and analysis of samples and the interpretation of results. But that applied work, if done well, still exhibits the same hallmarks of basic science: the use of validated methods and care in following their protocols; the development of careful and adequate documentation; the avoidance of biases; and interpretation conducted within the constraints of what the science will allow.

Adherence to scientific principles is important for concrete reasons: they enable the reliable inference of knowledge from uncertain information—exactly the challenge faced by forensic scientists. Thus, the reliability of forensic science methods is greatly enhanced when those principles are followed. National Research Council, *Strengthing Forensic Science in the United States: A Path Forward*, 111.

Washington law has grappled with evaluating the admissibility of scientific evidence. Essentially the There are three separate evaluations that must be made prior to admitting the results from a Datamaster breath machine. They involve a state statute, case law, and Rules of Evidence (ER 702 and ER 703).At first the three appear to be independent of each other. The statute insists a proper foundation is necessary while the case law and evidentiary rules insist on acceptance of the methodology, accuracy, and reliance of the data by the scientific community and that the proffered evidence will aid the trier of fact. Further analysis however, reveals all three share a universal requirement- that the underlying science and methodology supporting the evidence, in this case, a measurement, must be accepted by the scientific community as reliable and that the measurement is accurate and helpful to the trier of fact.

The evaluation dictated by statute applies exclusively to breath tests. In 2004 the legislature adopted a statute in response to legal challenges to the Datamaster. Essentially the RCW 46.61.506 4(b) requires admission of the breath test once the state lays the proper foundation (prima facie case) as it states:

> (b)For purposes of this section, "prima facie evidence" is evidence of sufficient circumstances that would support a logical and reasonable inference of the facts sought to be proved. In assessing whether there is sufficient evidence of the foundational facts, the court or administrative tribunal is to assume the truth of the prosecution's or department's evidence and all reasonable inferences from it in a light most favorable to the prosecution or department

However, scientific evidence offered to meet the prima facie standard must be based on a foundation of scientific facts and methodology and not speculation. "The 2004 amendment was intended to eliminate challenges to breath test admissibility based on technical deficiencies <u>not shown to adversely affect the accuracy of the result.(emphasis added)</u> *City of Fircrest v. Jensen,* 158 Wash.2d 384, 399, 143 P.3d 776 (2006). ,*Ludvigsen v. City of Seattle*, 162 Wash.2d 660, 681, 174 P.3d 43, 53 - 54(Wash.,2007). When the evidence of measurement relies on an inadequate scientific foundation, the proffered evidence must be classified as untrustworthy and inadmissible. To admit bogus and misleading science under a pretext of legitimate science is irrational and harmful to any notion of justice. Consequently the existence of substantial error affecting the scientific foundation of the evidence will result in a finding that the scientific foundation required has not been created by the Toxicology Lab.

The second method of evaluation, Washington case law, is consistent with the emphasis on following proper scientific principles as a precondition to admitting scientific evidence. Essentially, scientific evidence is admitted only when it is generally accepted to be reliable by the scientific community. In <u>State v. Copeland,</u> 130 Wash.2d 244, 922 P.2d 1304, 1312 (1996), the court said at 255:

> This court implicitly adopted the *Frye* standard for admissibility (cites omitted)… . The rationale of the *Frye* standard, which requires general acceptance in the relevant scientific community, is that expert testimony should be presented to the trier of fact only when the scientific community has accepted the reliability of the underlying principles. *Canaday,* 90 Wash.2d at 813, 585 P.2d 1185. "<u>In other words, scientists in the field must make the initial determination whether an experimental principle is reliable and accurate</u> *Id* (emphasis added)." The *Frye* standard recognizes that "judges do not have the expertise required to decide whether a challenged scientific theory is correct," and therefore courts "defer this judgment to scientists." *State v. Cauthron,* 120 Wash.2d 879, 887, 846 P.2d 502 (1993). … ."If there is a significant dispute between qualified experts as to the validity of scientific evidence, it may not be admitted." *Id.* at 887, 846 P.2d 502.

The Rules of Evidence constitute the third test to evaluate admission of scientific evidence. The purpose of introducing scientific evidence is to assist the trier of fact to understand the evidence or to determine a fact at issue. ER 702 requires that the witness qualify as an expert and that the testimony is helpful to the trier of fact. ER 703 addresses the foundational basis of opinion testimony. These two rules allow a trial court to suppress a breath test due to failures of accuracy or reliability of tests, or testing machines. *Jensen* 158 P2d. 395-9; *Ludvigsen* 174 P.3 at 54. Both evidentiary rules condition admissibility upon experts relying on the methodology employed by the scientists as a process that is accepted and relied upon by the scientific community, and for purposes other than litigation. State v. Nation, 110 Wn. App. 651, 451 P.3d. 1204 (2002).

While RCW 46.61506 (4)(b) requires a "proper foundation" and case law and the evidentiary rules follow *Frye* the universal requirement is that the evidence offered is based upon a scientific method or theory that is generally accepted in the scientific community; and that the evidence offered is accurate and reliable .

In summary as to the admission of Datamaster breath analysis, RCW 46.61. 506 at a minimum requires the evidence offered is grounded upon a foundation of scientific facts. *Frye* and ER 702 and ER703 all insist that the scientific evidence offered is complied in a manner that is acceptable and customarily employed by the scientific community.

ISSUE

IF SCIENTIFIC PRINCIPLES NECESSARY TO EVUALATE SCIENTIFIC EVIDENCE WHAT, SCIENTIFIC PRINCIPLES ARE APPLICABLE IN CASES INVOLVING BREATH TEST MEASUREMENTS?

In order to appreciate the scientific significance of the arguments of this case, an introduction to the applicable scientific principles presented is in order.
In 2006, The National Research Council was charged with examining and recommending improvements in the field of forensic science. In National Research Council, *Strengthing Forensic Science in the United States: A Path Forward*, 86 (2009) the authors wrote of the dynamic relationship between science and law at 86:

> Science and law always have had an uneasy alliance:
> Since as far back as the fourteenth century, scientific evidence has posed profound challenges for the law. At bottom, many of these challenges arise from fundamental differences between the legal and scientific processes.
> . . . The legal system embraces the adversary process to achieve "truth," for the ultimate purpose of attaining an authoritative, final, just, and socially acceptable resolution of disputes. Thus law is a normative pursuit that seeks to define how public and private relations *should* function. . . .
> In contrast to law's vision of truth, however, science embraces empirical analysis to discover truth as found in verifiable facts. Science is thus a descriptive pursuit, which does not define how the universe should be but rather describes how it actually *is*. (underlining added)

In order to create an accurate representation of the world, scientists employ the scientific method that has evolved over 800 years. There are two requirements that a forensic scientist must meet in order for his work to be reliable: satisfactory methodology and absence of error.

Evidence that applies forensic science bears the same scrutiny as test results produced in a laboratory. In criminal trials two questions must be addressed, National Research Council, *Strengthing Forensic Science in the United States: A Path Forward*, 87:

There are two very important questions that *should* underlie the law's admission of and reliance upon forensic evidence in criminal trials: (1) the extent to which a particular forensic discipline is founded on a reliable scientific methodology that gives it the capacity to accurately analyze evidence and report findings and (2) the extent to which practitioners in a particular forensic discipline rely on human interpretation that could be tainted by error, the threat of bias, or the absence of sound operational procedures and robust performance standards. These questions are significant.

The threat of bias is a particular concern in any measurement method, especially those methods using machines to measure. "Throughout scientific investigations the investigator must be as free from bias as possible, and practices are put in place to detect biases (such as those from measurements, human interpretation) and to minimize their effects on conclusions. National Research Council, *Strengthing Forensic Science in the United States: A Path Forward,* 112.(emphasis added).

There are standards that are recognized in the scientific community as universally acceptable for calibration laboratories. According to the testimony of Dr. Nayakk Pollister,KC7,p.221; and Dr. Ashley Emery, KC7,p.35 there are" principles that everyone should follow… the ISO people and (NIST and the ATSM), they have written them down and clarified them into standard practices". At KC7p.31-14,139 ,Dr. Emery noted that adherence to these standards is "necessary" to establish the credibility of scientific work and minimize error. These standards have not been met in the toxicology until November of 2009.

To summarize, scientific measurements are founded on two principles, the measurements must be based on a reliable methodology, and they must be accurate and free of bias as possible. Once these requirements have been met the court can consider the admissibility of the measurement. Properly measuring and employing tested methodology prevents the introduction of bogus science Forensic evidence that conflicts with applicable scientific principles is too unreliable to be considered by the court. This evidence is not admissible.

ISSUE

BY OMITTING BIAS AND UNCERTAINITY IN MACHINE MEASUREMENTS, OF REFERENCE MATERIALS, HAS THE WASHINGTION STATE TOXOLOGY LAB COMPROMISED THE DATAMASTER'S ACCURACY SO AS TO RENDER THE MACHINE'S OUTPUT INADMISSIBLE BEFORE THE TRIER OF FACT?

The burden of proof lies with the state. *State v. Baker,* 56Wn.2d 846,355P2d 806 (1960). Consequently the state must request for admission of the breath test must meet scientific principles.

"We need to acknowledge there's uncertainty in every measurement and that bias is part of that uncertainty." Gullelnburg, KC2,p.78.

Bias is a systematic error in a measurement. The error is systematic because the machine will consistently under or overestimate a true value by a certain amount.

According to Rod Gullenburg "There is uncertainty in every measurement... virtually always some degree of bias, but that's ok as long as.... .it can be corrected for." Gullenburg KC2, p.78. The Toxicology director admitted that no adjustment is made for bias in producing any simulator solutions. Logan, KC5,p.10-1.

Uncertainty can result from of human error, or some type of bias other than systematic error. In any case, bias and uncertainty must be accounted for. Bias is a necessary adjustment to a measurement of breath. In testing, the scientific community requires adjustment for bias. "Throughout scientific investigations the investigator must be as free from bias as possible, and practices are put in place to detect biases (such as those from measurements, human interpretation) and to minimize their effects on conclusions. National Research Council, *Strengthing Forensic Science in the United States: A Path Forward*, 112.

In order to calibrate the Datamaster, a reference solution containing specific levels of alcohol and water is made by the Toxicology lab. The toxicology lab measures the concentrations water and alcohol of the solution by employing a gas chromatograph. However bias has not been measured on this machine as to the concentration of the simulator solution (QAP solution), nor were any adjustments made for other types of bias such as rounding or recording or calculation errors.

Failing to consider bias and uncertainty is important because the simulator solution produced by the Toxicology Lab is the only means to calibrate the Datamaster. A faulty QAP solution will result in an erroneous calibration of the Datamaster.

For evidentiary considerations, bias is a critical component to determining whether the methodology is accepted as customary and reliable in the scientific community. The state's witness in the King county case of State v. Ahmach et al.C00627921 was Mr. Rod Gullberg. He testified that we must start with an accurate reference value for our simulator (QAP or field) solution. I need to know (bias) because I am relying on the reference value from that gas chromatograph to identify this simulator solution concentration. Gullelnburg,KC2,p66-7

In the instant case the omission of calculating for bias and uncertainty permeates the development and use of the reference solutions both for the QAP solution and the field test solutions. The entire Datamaster process depends on the accuracy of the solutions used to calibrate individual Datamasters. Due to neglecting bias in measuring gas chromatographs, and failing to adjust for uncertainty, the concentration of any reference solution is not scientifically reliable or accurate.

Another omission of significance is the Toxicology Lab's practice of discarding outliers. An outlier is an observation that is numerically distant from other readings in the sample measured. This policy leads to underestimating measurements and introduces uncertainty and a lack of precision in measured values. INTERNATIONAL ORGANIZATION FOR STANDARDIZATION ,Reference Materials- General and Statistical; Principles for Certification, ISO Guide 35:2006Esec 10.5.5 (2006).

Discarding outliers creates the risk of never finding anything unanticipated and must be resisted strongly. Gullberg, KC2, p.15.

The Toxicology Lab's errors and neglect to follow accepted scientific practices are substantial. Typical errors include: misreporting data, mixing up certifications so that chromatograms and solutions did not match. Logan,KC5 p.99-100, Gullberg, KC1,p.127-8. "Forensically indefensible" calculations of vapor concentrations. Gullberg, Kc1,p.117,120. Inclusion of bad data caused by a malfunction of the chromatograph. Dr. Nuwayhid KC6,p.164-5. Computer programming errors which yielded incorrect Equivalent Vapor Concertrations (EVC), thereby making it harder for a defendant to determine bias on the Datamaster. Gullberg, KC2,p.64,77, 107-8,213-4.

The cumulative effect of the violations of scientific principles renders the measurement of breath unreliable, and inaccurate. The methodology injects far too much carelessness and uncertainty into the scientific process that produces a breath analysis by the Datamaster. The breath analysis proffered by the Toxicology Lab is missing a scientific foundation necessary to present a prima facie case required by RCW 46.61.506 (4)(a)(vi). Specifically the simulator solution (whether a Field or QAP solution) used to provide an known alcohol concentration to calibrate the Datamaster is <u>not scientifically reliable</u> as required by WAC 448-16-030(8). The solution is not a known standard that can be a basis for determining accuracy. Nor could there be peer review. The breath analysis could not be repeated because there is not bias reading to recreate the solution or to measure it by third parties.

The *Frye* standard to evaluate admitting breath test results requires the scientific methodology employed by the proponent to be generally accepted in the scientific community. The numerous errors and carelessness identified are so persuasive that the scientific community would not accept such a performance. Consequently applying the *Frye* test would render the breath analysis inadmissible.

Even if *Frye* was satisfied, breath measurements would not clear the scrutiny of ER 702 and ER703. To be admissible under ER 702 the witness must be an expert and the testimony is helpful to the trier of fact. The expertise is questionable given the score of errors witnessed in its operation. Presenting a breath ticket based on a reference solution that has not even been adjusted for bias and uncertainty would introduce misleading information to the trier of fact. In short if the breath test is unreliable the evidence would not be helpful to the trier of fact. The flawed evidence would be inadmissible. *Cauthron*, 120 Wn.2d at 890.

ER703 allows evidence that is of the type that is relied upon by experts in the field. Again the omission of outliers, bias, and uncertainty, when measuring reference solutions alone is sufficient to conclude that a professional scientist would not rely on the work product of the Toxicology' Lab. The lack of scientific reliance on the breath measurement renders the toxicology Lab's breath analysis inadmissible.

CONCLUSION

The litany of errors and omissions exhibited by the Toxicology Lab are numerous and substantial. The cumulative effect violates accepted standards of reliability and accuracy required by the scientific community. The breath measurement is not grounded on a firm scientific foundation and cannot be considered to be reliable, accurate, or helpful to a trier of fact. The Toxicology Lab has not demonstrated adherence to accepted scientific principles, consequently the breath measurement is inaccurate and inadmissible.

DEFENDANT'S MOTION TO SUPRESS IS GRANTED.
THIS ORDER WILL REMAIN IN EFFECT UNTIIL THE TOXICOLOGY LAB METHODOLOGY AND OPERTION IS CERTIFIED TO ISO STANDARDS.

Patrick O'Malley, Judge 3 December 2009

APPENDIX D

TED VOSK - CV

TED VOSK
8105 NE 140th Pl.
Bothell, WA 98011
P: (425) 753-6343 ◊ F: (425) 820-7532
tvosk@comcast.net

EDUCATION

UNIVERSITY OF WASHINGTON
- Information Assurance and Cybersecurity program. 2006
- Graduate courses in physics. 2001 – 2003

HARVARD LAW SCHOOL 1999
Juris Doctor
Thesis: *Human Cloning and FDA Regulation*.
- Environmental Law Review.
- Hale and Dorr Legal Services Center.
- Harvard Defenders.

CORNELL UNIVERSITY 1995 – 1996
Graduate (Ph.D.) studies in Physics.

EASTERN MICHIGAN UNIVERSITY 1995
Bachelor of Science, Physics and Mathematics
Magna cum laude
Honors Theoretical Physics
Thesis: *A Comparative Spectroscopic and Topographic Analysis of the Surface of Graphite.*
- University Honors College.
- Phi Kappa Phi.
- Sigma Pi Sigma.
- Golden Key National Honor Society.
- The Stoic Society: *Honor Fraternity of EMU*.
- President, Society of Physics Students: *EMU Chapter*.
- Editor/Writer: *E.M.U. Astronomy Club Newsletter*.

UNIVERSITY OF MICHIGAN, SPACE PHYSICS RESEARCH LAB Summer, 1993
- *Research Associate*: Computer modeling of atomic oxygen concentration and dynamics in the mesosphere and thermosphere based on rocket-collected data.

FORENSICS TRAINING

GOOD MEASUREMENT PRACTICES IN THE PROPER USE AND CALIBRATION OF BALANCES AND PIPETTES Seattle, WA
AMERICAN ACADEMY OF FORENSIC SCIENCES - WORKSHOP Feb., 2010
Faculty: *Dr. Thomas Brettell, Dept. of Chemical & Physical Sciences, Cedar Crest College*
Joe Moran, Metrology Manager for Troemner Calibration Technologies & Precision Weights
Janine Kishbaugh, Quality Control Manager Forensic Sciences, Cedar Crest College

DEPAUL COLLEGE OF LAW — Chicago, Ill
CENTER FOR LAW AND SCIENCE — May 7-9, 2009
SCIENCE IN THE COURTROOM FOR THE 21ST CENTURY: ISSUES IN FORENSIC DNA
Faculty: *Dr. Dane Krane, Biological Sciences, Wright State University*
Dr. Jason Gilder, Forensic Bioinformatics Inc.
Dr. Norah Rudin, Forensic DNA Consulting

ISO/IEC 17025:2005 SECTION 5.4.6: ESTIMATION OF UNCERTAINTY — Denver, CO
AMERICAN ACADEMY OF FORENSIC SCIENCES - WORKSHOP — Feb., 2009
Faculty: *Joseph Bono, Laboratory Director, United States Secret Service, Forensic Services Division*
Dr. Elizabeth Mishalanie, EPA National Enforcement Investigations Center

ALCOHOL TESTING: CRITICAL IMPORTANCE OF METHOD VALIDATION — Phoenix, AZ
Faculty: *Dr. A.W. Jones, Nat'l. Board of Forensic Medicine, Swedish Government* — May, 2008

THE AGORA – FALL WORKSHOP — Seattle, WA
- PROJECT 610 (An Infosec Penetration of the Seattle Police Department) — Sept., 2006
- U.S. V. CHRISTOPHER MAXWELL: THE INVESTIGATION AND CAPTURE OF A BOTNET OPERATOR

Faculty: *Kathryn Warma, Assistant United States Attorney*
Dave Farquhar, FBI Special Agent

INTOXILYZER 8000 OPERATOR'S COURSE — New Orleans, LA
Faculty: *Dr. Alfred Staubus, College of Pharmacy (Pharmaceutics), Ohio State University* — Sept., 2006

NHTSA/IACP STANDARDIZED FIELD SOBRIETY TESTING INSTRUCTOR COURSE — San Antonio, TX
WALDEN, PLATT & ASSOCIATES: IMPAIRED DRIVING CONSULTING — June, 2005

NHTSA STANDARDIZED FIELD SOBRIETY TESTING PRACTITIONER COURSE — Seattle, WA
WALDEN, PLATT & ASSOCIATES: IMPAIRED DRIVING CONSULTING — May, 2005

DRUG EVALUATION AND CLASSIFICATION (DECP) OVERVIEW COURSE — Seattle, WA
WALDEN, PLATT & ASSOCIATES: IMPAIRED DRIVING CONSULTING — Oct., 2004

THE PHYSIOLOGY OF BREATH ALCOHOL TESTING — Las Vegas, NV
Faculty: *Dr. Michael Hlastala, Physiology, Biophysics and Medicine, Univ. of Washington* — March, 2004

TEACHING ACTIVITIES – UNIVERSITY RELATED

UNIVERSITY OF WASHINGTON — Seattle, WA
EVANS SCHOOL OF PUBLIC AFFAIRS — May, 2009
Guest Lecturer
- Topics in Science, Technology, and Public Policy: Policy Formulation and Implementation

UNIVERSITY OF WASHINGTON — Seattle, WA
SCHOOL OF LAW — Oct., 2008
Judge
- 1L Mock Trial Competition

EDMUNDS COMMUNITY COLLEGE — Edmunds, WA
Guest Lecturer — 2007 – 2008
- Business Law: BUS 240

EASTERN MICHIGAN UNIVERSITY — Ypsilanti, MI
COLLEGE OF ARTS & SCIENCES SYMPOSIUM XXV — April, 2005
Moderator
- Session on physics and astronomy.

CORNELL UNIVERSITY — Ithaca, NY
Graduate Teaching Associate — 1995 – 1996
- PHY 330: Modern Experimental Optics.
- PHY 101 & 102: General Physics I & II.

EASTERN MICHIGAN UNIVERSITY — Ypsilanti, MI
Teaching Assistant — 1993 – 1994
- PHY 223 & 224: Introductory physics labs.

NON-PROFIT / PUBLIC EDUCATION

CELESTIAL NORTH — Seattle, WA
Vice President, Writer, Broadcaster, Public Speaker — 2003 – 2008
- Washington 501(c)(3) focusing on education in astronomy and space sciences for K-12 and lay public.
- "It's Over Your Head" radio program broadcast on KSER 90.7 FM: Writer & on air personality.

EMPLOYMENT

ATTORNEY AT LAW/CONSULTANT/LEGAL & SCIENCE WRITER — Seattle, WA
Criminal Defense/Appeals/Administrative Law/Forensic Metrology — 2004 – Present
Of Counsel, COWAN, KIRK, GASTON (2009 – present)
Of Counsel, CALLAHAN LAW (2007 – 2009)

MAGNUSON LOWELL — Redmond, WA
Attorney, Criminal Defense/Tort/Administrative Law — 2003 – 2004

UNIVERSITY OF WASHINGTON — Seattle, WA
DEPARTMENT OF CHEMISTRY
- NSF SCIENCE & TECHNOLOGY CENTER FOR INFORMATION TECHNOLOGY RESEARCH — 2002 – 2003
  Acting Managing Director
- DALTON RESEARCH GROUP — 2001 – 2003
  Research Program Manager

TUCKER & STEIN — Bellevue, WA
Public Defender — 2001

CITY OF REDMOND PROSECUTOR'S OFFICE — Redmond, WA
Deputy Prosecutor — 2000 – 2001

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 OREGON (*Pro Hac Vice* – Trial and Appellate counsel) 2006 – 2008

NOTABLE CASES

- *City of Seattle v. Winebrenner*, 219 P.3d 686 (Wash. 2009): Amicus Curiae – Wash. Assoc. of Crim. Def. Lawyers.
- *Washington State Toxicology Lab Litigation*: Lead counsel litigation exposing systemic misconduct, carelessness and incompetence in the Washington State Toxicology Lab. Resulted in resignation of several officials including State Toxicologist, suppression of evidence around state and reform of the Lab.
  - *State v. Ahmach*, <http://www.metrokc.gov/kcdc/ordgrbac.pdf> (2008).
  - *Arntson v. Dept. of Licensing*, <www.waduicenter.com/toolbox/documents/ArntsonRuling.pdf> (2007).
- *Ludvigsen v. City of Seattle*, 174 P.3d 43 (Wash. 2007).
- *Butler v. Kato*, 154 P.3d 259 (Wash. App. 2007): Consultant.
- *City of Fircrest v. Jensen*, 143 P.3d 776 (Wash. 2006).
- *Devine v. Dept. of Licensing*, 110 P.3d 237 (Wash. App. 2005): Consultant.
- *City of Seattle v. Clark-Munoz*, 93 P.3d 141 (Wash. 2004): Consultant.
- *City of Bellevue v. Tinoco*, < http://your.kingcounty.gov/KCDC/BELTHERM.PDF > (2001).

PROFESSIONAL MEMBERSHIPS

- AMERICAN ACADEMY OF FORENSIC SCIENCES (AAFS)
- NATIONAL ASSOCIATION OF CRIMINAL DEFENSE LAWYERS (NACDL)
- WASHINGTON ASSOCIATION OF CRIMINAL DEFENSE LAWYERS (WACDL)
- NATIONAL COLLEGE FOR DUI DEFENSE (NCDD)
- WASHINGTON FOUNDATION FOR CRIMINAL JUSTICE (WFCJ)
- AMERICAN ASSOCIATION FOR THE ADVANCEMENT OF SCIENCE (AAAS)
- AMERICAN SOCIETY FOR TESTING AND MATERIALS INTERNATIONAL (ASTM)
- AMERICAN SOCIETY FOR QUALITY (ASQ)
- AMERICAN PHYSICAL SOCIETY (APS)
- AMERICAN CHEMICAL SOCIETY (ACS)
- AMERICAN MATHEMATICAL SOCIETY (AMS)
- MATHEMATICAL ASSOCIATION OF AMERICA (MAA)
- ASSOCIATION FOR SYMBOLIC LOGIC (ASL)

Honors & Awards

- ❖ American Academy of Forensic Sciences (Assoc. member) 2009
- ❖ President's Award, *Washington Association of Criminal Defense Lawyers.* 2008
- ❖ Pro Bono Public Service Commendation, *Washington State Bar Association.* 2008
- ❖ Certificate of Distinction, *Washington Foundation for Criminal Justice.* 2007
- ❖ Pro Bono Public Service Commendation, *Washington State Bar Association.* 2007
- ❖ Profiled in Harvard Law Bulletin: *Celestial Reasonings*, Harv. L. Bull., Spring 2007, at 52. 2007
- ❖ Pro Bono Public Service Commendation, *Washington State Bar Association.* 2006
- ❖ Out of This World Award for Excellence in Astronomy Outreach, *Astronomy Magazine.* 2006
- ❖ Super Lawyer Rising Star, *Washington Law & Politics Magazine.* 2005
- ❖ Mensa. 2001
- ❖ Goldwater Scholar in Mathematics, Science and Engineering, *Goldwater Foundation*. 1993 – 1995
- ❖ Outstanding Student of Mathematics: *E.M.U. Dept. of Mathematics.* 1995
- ❖ Lobbestael Scholarship - Achievement in Mathematics: *E.M.U. Dept. of Mathematics.* 1993, 1994
- ❖ Leib Scholarship - Superior Performance in Physics: *E.M.U. Dept. of Physics.* 1992, 1993, 1994
- ❖ Recognition of Excellence Scholarship: *E.M.U. College of Arts & Sciences.* 1992, 1993, 1994
- ❖ Campus Leader Scholarship: *E.M.U. College of Arts & Sciences.* 1994
- ❖ Robert Silver Award - Outstanding Scholarship in Physics: *E.M.U. Dept. of Physics.* 1993

Political/Legislative Activities

Washington Association of Criminal Defense Lawyers — Seattle, WA
Legislative Committee Member — 2006 – Present
Forensic Investigations Council Legislation Workgroup — 2009 – Present

Jeanette Dalton Judicial Campaign Committee — Port Orchard, WA
Communications Consultant — 2008

Democratic Party, Voter Protection Program — Seattle, WA
Election Monitor — 2004, 2006, 2008

Publications

Texts/Treatises

- ❖ Forensic Metrology: A Primer on Scientific Measurement for Lawyers, Judges and Forensic Scientists (CRC Press, *In Preparation*).

- ❖ Understanding DUI Scientific Evidence, 3rd ed. (Aspatore, *In Press*).
  - Chapter __: DUI Evidence and the National Academy of Sciences' Report on Forensic Science.

- ❖ Intoxication Test Evidence, 2nd ed. (Thomson-West, 2009).
  - Chapter 56: Uncertainty in Forensic Breath Alcohol Testing.

- ❖ Understanding DUI Scientific Evidence, 2nd ed. (Aspatore, 2009).
  - Chapter 3: Field Sobriety Testing and Driver Impairment: Linked or Not?

- ❖ Washington DUI Practice Manual (Wash Prac. Series, v.32)(Thomson-West, 2008).
  - Chapter 21: Standardized Field Sobriety Testing.
  - Appendix: Toxicology Lab Argument Summary.

- ❖ Defending DUI's in Washington, 3rd ed. (LexisNexis 2008).
  - Chapter 13: The DataMaster.

- Chapter 15: DIRECT EXAMINATION OF THE DEFENSE EXPERT.

PERIODICALS

- *Chaos Reigning: Breath Testing and the Washington State Toxicology Lab*, THE NACDL CHAMPION, June 2008.
- *Down the Rabbit Hole: The Arbitrary World of the Washington State Toxicology Lab*, WASH. CRIM. DEF., May 2008.
- *Due Process and Science by Legislative Decree*, WASH. CRIM. DEF., Feb. 2007.
- *Precluding Standardized Field Sobriety Tests in non-per se Prosecutions*, WASH. CRIM. DEF., Feb. 2006.

CONTRACT ARTICLES

- *Field Sobriety Tests: Another Government Lie?*, BAR NEWS, August 2007.
- *A New Paradigm for Challenging Breath Test Evidence in Washington*, BAR NEWS, June 2007.

PRESENTATIONS

AAFS 2010 ANNUAL MEETING — Seattle, WA
AMERICAN ACADEMY OF FORENSIC SCIENCES — Feb., 2010

- *Workshop Co-chair & Faculty*: Attorneys and Scientists in the Courtroom: Bridging the Gap.
  - *Faculty*: Metrology: A Knowledge Base for Communication and Understanding.
- *Session Moderator*: DNA II

DISCOVERY: IT'S ELEMENTARY — Bellevue, WA
WASHINGTON ASSOCIATION OF CRIMINAL DEFENSE LAWYERS — Feb., 2010

- *Faculty*: Fun With Bodily Fluids: How to Make the Toxicology Lab Your Best Witness.

DEFENDING DUIS — SeaTac, WA
WASHINGTON FOUNDATION FOR CRIMINAL JUSTICE — Dec., 2009

- *Faculty*: Breath Testing – Beating the Odds.

DEFENDING DUIS — Spokane, WA
SPOKANE COUNTY PUBLIC DEFENDERS OFFICE — Sept., 2009

- *Faculty*: Defending DUIs.

FORENSIC METROLOGY: THE BASICS — Seattle, WA
WASHINGTON DEFENDER ASSOCIATION — Sept., 2009

- *Faculty*: Forensic Metrology: The Basics

DEPAUL COLLEGE OF LAW — Chicago, Ill
CENTER FOR LAW AND SCIENCE — May, 2009
SCIENCE IN THE COURTROOM FOR THE 21ST CENTURY: ISSUES IN FORENSIC DNA

- *Faculty*: Forensic Metrology – Why it's Essential to Litigating Forensic Science Cases.
- *Panelist*: The Implications and Opportunities of the National Academy of Sciences Report on Forensic Science.

NATIONAL FORENSIC BLOOD AND URINE TESTING SEMINAR — San Diego, CA
GEORGIA ASSOCIATION OF CRIMINAL DEFENSE LAWYERS — May, 2009

- *Faculty*: Forensic Metrology: The Key to the Kingdom.
- *Panelist*: Decisions From Around the USA to Help You Win Tough Cases.

CARS AND CRIME: ELECTRIFYING YOUR DEFENSE — Seattle, WA
WASHINGTON ASSOCIATION OF CRIMINAL DEFENSE LAWYERS — March, 2009
❖ *Faculty*: The National Academy of Science's Report and the Future of Forensic Science in Washington.

DEFENDING DUIS — Everett, WA
SNOHOMISH COUNTY BAR ASSOCIATION — Nov., 2008
❖ *Faculty*: Are Breath Test Results Accurate?

ANNUAL CONFERENCE — Tacoma, WA
WASHINGTON COURT REPORTERS — Oct., 2008
❖ *Speaker*: Washington State Toxicology Lab Controversy.

THE 15TH ANNUAL CRIMINAL JUSTICE INSTITUTE — Seattle, WA
WASHINGTON STATE BAR ASSOCIATION — Sept., 2008
❖ *Faculty*: Toxicology Lab Results — Where do we go from here?

ANNUAL CONFERENCE: EQUAL JUSTICE FOR ~~ALL~~ SOME? — Chelan, WA
WASHINGTON ASSOCIATION OF CRIMINAL DEFENSE LAWYERS — June, 2008
❖ *Faculty*: DUI/Crime Lab Update.

SCIENCE AND THE LAW — Seattle, WA
NORTHWEST DEFENDERS ASSOCIATION — June, 2008
❖ *Faculty*: Scientific Standards – How to Master Science Without Being a Scientist.

DEFENDING DUIS — SeaTac, WA
WASHINGTON FOUNDATION FOR CRIMINAL JUSTICE — Dec., 2007
❖ *Faculty*: Motions – Crippling the State's Case.

DISTRICT COURT PRACTICE — Tacoma, WA
WASHINGTON ASSOCIATION OF CRIMINAL DEFENSE LAWYERS — Oct., 2007
❖ *Faculty*: Challenging the Crime Lab.

BASICS OF DUI DEFENSE: FROM THE LAB TO THE COURTROOM — Seattle, WA
WASHINGTON DEFENDER ASSOCIATION — Oct., 2007
❖ *Faculty*: The State Toxicology Lab.

STATE TOXICOLOGY LAB: WHAT DEFENDERS NEED TO KNOW — Seattle, WA
WASHINGTON DEFENDER ASSOCIATION — Sept., 2007
❖ *Faculty*: Breath Tests and the State Toxicology Lab.

DUIS: FROM TRAFFIC STOP TO TRIAL — Seattle, WA
SPOKANE COUNTY BAR ASSOCIATION — Sept., 2007
❖ *Faculty*: The Washington State Toxicology Lab – What you need to know.

DUIS IN A POST-JENSEN WORLD — Seattle, WA
WASHINGTON DEFENDER ASSOCIATION — Feb., 2007
❖ *Faculty*: Techniques, Strategies and Methods for Addressing Breath Tests.

ATLA 2006 Annual Convention — Seattle, WA
DUI Defense Tips From the Masters — July, 2006
Association of Trial Lawyers of America
- *Co-author*: Field Sobriety Tests: "Tell Them No!"

Defending DUIs — Bellevue, WA
Washington State Office of Public defense — March, 2006
- *Faculty*: The Mathematical Analysis of Breath Testing.
- *Faculty*: Standardized Field Sobriety Testing.

Table Mountain Star Party — Ellensburg, WA
Table Mountain Star Party Association, Ltd. — August, 2005
- *Speaker:* Strings, Gravity, and Locality: An Overview of Modern Cosmology.

National College for DUI Defense, Summer Session 2005 — Cambridge, MA
National College for DUI Defense — July, 2005
- *Co-author/Editor*: The Bad-Bad Case and Blood Tests (Volume 2 of text).

Undergraduate Research Symposium XXV — Ypsilanti, MI
Eastern Michigan University College of Arts and Sciences — April, 2005
- *Moderator*: Section on Physics and Astronomy.

A Constitutional Challenge to the 2004 DUI Law — Seattle, WA
Washington Alliance of DUI Lawyers — Dec., 2004
- *Faculty*: Constitutional Separation of Powers.

Table Mountain Star Party — Ellensburg, WA
Table Mountain Star Party Association, Ltd. — August, 2004
- *Speaker:* Celestial North: Awe, Wonder and the Need to Know.

Undergraduate Research Symposium XV — Ypsilanti, MI
Eastern Michigan University College of Arts and Sciences — June, 1995
- *Research Presentation:* Site Asymmetry in Scanning Tunneling Microscopy of Graphite.

National Undergraduate Research Symposium — Argonne, IL
Argonne National Laboratory — Nov., 1994
- *Research Presentation:* Site Asymmetry in Scanning Tunneling Microscopy of Graphite.

Undergraduate Research Symposium XIII — Ypsilanti, MI
Eastern Michigan University College of Arts and Sciences — June, 1993
- *Research Presentation:* Communication via Laser Beam.

Undergraduate Research Symposium XII — Ypsilanti, MI
Eastern Michigan University College of Arts and Sciences — June, 1992
- *Research Presentation:* Photoelectric Photometry of Small Amplitude Red Variables.

Television/Radio Appearences

- The Dave Ross Show: 710 KIRO (8/6/08).
- Up Front with Robert Mak: KING 5 (2/2/08).

Miscellaneous Activities

- SCUBA certifications:
  - Rescue diver.
  - Advanced open water diver.
- Capital City Marathon, 2003.
- Motorcycling.
- Sky Diving.
- White Water Rafting.
- Skiing/Snow-boarding.
- Surfing.
- Astronomy.