IN THE CIRCUIT COURT,
NINTH JUDICIAL CIRCUIT,
IN AND FOR ORANGE
COUNTY, FLORIDA.

CASE NO.: 48-2008-CF-15606

DIVISION: 99

WITH ATTACHMENTS

FILED IN OFFICE
CRIMINAL DIVISION
2010 DEC 29 PM 5:09
LYDIA GARDNER
CLERK CIRCUIT COURT
ORANGE CO., FL

STATE OF FLORIDA

VS.

CASEY MARIE ANTHONY
__________________________/

MOTION TO EXCLUDE UNRELIABLE EVIDENCE

Defendant, Casey Marie Anthony, by and through her undersigned attorneys, moves this Honorable Court to exclude from this cause any testimony or evidence concerning any alleged identification of the chemical composition of human decomposition odor, any testimony regarding a test involving elemental analysis of Laser Induced Breakdown Spectroscopy, any testimony regarding quantification of chloroform, or reference to an alleged "decompositional" odor analysis database pursuant to Sections 90.401 and 90.402 and 90.403, Florida Statutes, the due process clauses of Amendments Five and Fourteen, Constitution of the United States and Article 1, Section 9, Florida Constitution, and *U.S. v. Frye*, 293 F. 1013 (D.C.Cir. 1923). As general legal grounds, Defendant states:

1. Florida recognizes a test first enunciated in *U.S. v. Frye*, 293 F. 1013 (D.C. Cir. 1923). This test is imposed as a threshold for admissibility of a scientific principle or test. Under Frye, it must be shown that a scientific principle or test is "sufficiently

established to have gained general acceptance in the particular field in which it belongs." Id., 293 F. at 1014. This ensures a jury will not be mislead by experimental scientific methods that may ultimately prove to be unsound. *See Stokes v. State*, 549 So.2d 188 (Fla. 1989). *Stokes* holds that a **"courtroom is not a laboratory, and as such it is not the place to conduct scientific experiments. If the scientific community considers a procedure or process unreliable for its own purposes, then the procedure must be considered less reliable for courtroom use."**

2. As outlined by the Florida Supreme Court in *Ramirez v. State*, 651 So.2d 1164 (Fla. 1995), the *Frye* test is a four step process. The steps are: 1) whether such expert testimony will assist the jury in understanding the evidence or in determining a fact in issue, 2) whether the expert's testimony is based on a scientific principle or discovery that is sufficiently established to have gained general acceptance in the particular field in which it belongs, 3) whether a particular witness is qualified as an expert to present opinion testimony on the subject in issue, and 4) the jury's determination of the credibility of the expert opinion, which it may either accept or reject. Under *Ramirez*, it is up to the proffering party to demonstrate the requirements of both scientific reliability and general acceptance in the field.

3. "Pure opinion" testimony is not subject to the *Frye* test. Pure opinion testimony is testimony which "does not rely upon any study, test, procedure or methodology that constitutes new or novel scientific evidence." *Gelsthorpe v. Weinstein*, 897 So.2d 504. at 510-511(Fla. 2d DCA 2005), quoted with approval in *Marsh v. Valyou, Jr.* 977 So.2d 543 (Fla. 2007). The evidence which Defendant seeks to exclude

by this motion does not constitute pure opinion testimony. Additionally said testimony is not deduced from a well-recognized or established scientific methodology.

4. Introduction of testimony concerning "test" results when the scientific community does not consider the methodology used to deduce those results to be reliable and which do not meet the *Frye* test would result in a denial of due process under both the Florida and Federal Constitutions because admission of such testimony or evidence would lead to jury confusion about which evidence is reliable and which is unreliable. Due process of law is a constitutional guarantee of respect for personal rights that is "so rooted in the traditions and conscience of our people as to be ranked as fundamental." *Snyder v. Massachusetts*, 291 U.S. 97, 105 (1933). A jury verdict premised upon testimony or evidence which should not have been admitted because it embraces a scientific principle or methodology which is unreliable will violate Defendant's due process rights enunciated in Article 1, Section 9 of the Florida Constitution and the Fifth and Fourteenth Amendments of the United States Constitution.

5. Additionally, introduction of testimony concerning scientific principles or tests that the scientific community does not consider reliable or testimony deduced from an unreliable methodology which do not meet the *Frye* test would violate Sections 90.401 and 90, 402, Florida Statutes as immaterial to the facts at issue and violate Section 90.403, Florida Statutes, in that the probative value of such evidence is greatly outweighed by the prejudicial effect of such evidence.

A statement of the facts that the State seeks to introduce relating to decomposition odor is as follows:

1. Dr. Vass is listed as a prosecution witness. He is an alleged "research scientist" with the Oak Ridge National Laboratory, Oak Ridge, Tenn. He purports to have been conducting research since 2002 in an attempt to identify the chemical composition of human decomposition odor and claims to have developed a "decompositional odor analysis database."

2. In his report, Dr. Vass stated that he identified 54 chemicals from a sealed metal can and identified 54 chemicals from the "concentrated" sample. Of these, 43 (79.6%) are consistent with "decompositional events". Certain fluorinated compounds which he claims are usually associated with human decomposition **were not** detected. He suggests a possible reason for the absence of these findings but acknowledges this reason has not been "studied". He acknowledged that his findings "do not rule out the possibility that an animal carcass, rotting meat, paint varnish cleaners degreasers or garbage were transported in the trunk at some time that may have contributed to the observed chemical compounds."

3. At deposition Dr. Vass admitted that he is a research scientist and that his laboratory is not a forensic laboratory, therefore he is not required to have any of the following:

a. Standard Protocols: This is defined as a written standard in which tests are to be performed, including a checklist of steps to be taken to ensure that the test is repeatable and valid.

b. Quality assurance: This is defined as quality control, a list of regulations or procedures to ensure that not only are the examiners sufficiently trained, but this is a common method to catch and limit contamination and the integrity of results.

c. Error rates: This is defined as a statistical calculation of confirmed results and errors.

He further testified that he is unaware of any other individual in the entire country who uses his "methodology" for determining human decomposition. Dr. Vass, through his counsel, has refused to disclose the underlying data to support his conclusions and in essence is taking a "trust me" approach to science.

6. At deposition, Dr. Vass admitted that he did not collect the sample or personally inspect the car from which the sample was collected. He was unconcerned about the circumstances concerning the collection of the sample including facts that the car was located in a tow yard, a garage at the Anthony home, and law enforcement garage, the contents of the garage could not be ascertained at the time the sample was taken and the trunk had been opened to the garage air for many hours. There is no standard protocol for the collection of the sample.

7. The "test" performed by Dr. Vass cannot be duplicated or verified by any other expert in the world because no other scientist does this type of experiment or study. The air sample is impossible to preserve at the time of the collection of evidence and even more so to the time Miss Anthony was last seen driving the vehicle. Furthermore, there are no protocols either "accepted" or "unaccepted" for duplication of the test, this includes the Oakridge laboratory. Thus, the defense has been unable to have a defense expert review the testing and makes it impossible for an independent expert to test and review the work done.

8. Dr. Vass' work has never been validated in a Court of law anywhere in the world. His opinions and theory's range from the interesting to the bizarre, Dr. Vass is

attempting to develop a mobile sniffer machine to find clandestine graves, as well as put electronic leashes on flies. While interesting and well intended this is exactly the science that Frye and its progeny preclude.

OAKRIDGE LABORATORIES "SECRET SCIENCE" VIOLATES MISS ANTHONY'S RIGHT TO CONFRONTATION

Upon receipt of Dr. Vass' "forensic report," the Defense requested directly from Oak Ridge National laboratories a complete list of all 424 chemical compounds that Dr. Vass believes consists of human decomposition.[1] The Oak Ridge National Laboratories by and through their counsel Mr. Allen Parker and Assistant State Attorney Jeffery Ashton acting as a "facilitator" responded to our request by stating that the chemical database and it's information is proprietary to the agency that funded the research (Through deposition it was discovered that this agency is the FBI). Therefore, the Oakridge National Laboratories refused to comply with this discovery request and is choosing to mask the empirical data relied upon for it's conclusions and has taken a rare unheard of "trust me" approach to science. Miss Anthony filed a motion to compel this information with this Honorable Court. On June 1st 2010 that request was denied without prejudice.

The Six Amendment of the United States allows one to confront their accusers, it is through this concept of law that permit such liberal discovery in the State of Florida. Without Dr. Vass' "secret formula" it is impossible for Miss Anthony to:

a. Conduct and independent examination of the evidence or theories submitted by Dr. Vass.

---

[1] Letter dated July 22, 2009 to Oakridge National Laboratories from Linda Kenney Baden. (previously submitted)

b. Properly and effectively cross examine Dr. Vass when testifying at the Frye hearing before this Honorable Court.

Thus it is the position of the Defense that to deny access to Dr. Vass' database of human decomposition denies Miss Anthony's right to confrontation and would renew it's objection and move this Honorable Court to compel Oakridge National Laboratories all discovery items to the defense prior to any Frye hearing conducted by this Honorable Court.

TESTIMONY REGARDING HUMAN DECOMPOSITION ODOR AS IT RELATES TO THIS CASE IS INHERENTLY UNRELIABLE AND THE RESULT OF THE APPLICATION OF PRINCIPLES THE SCIENTIFIC COMMUNITY DOES NOT CONSIDER RELIABLE AND WHICH DO NOT MEET THE *FRYE* TEST

In *Flanagan v. State*, 586 So.2d 1085 (Fla. 1st DCA 1991), reversed, 625 So.2d 827 (Fla. 1993), the Florida Supreme Court considered whether sex offender profile evidence met the *Frye* test for admissibility. At issue was whether the testimony of an HRS Child Protection Team psychologist who testified about "common characteristics of the home environment where child sexual abuse occurs and about the characteristics of abusers" was admissible as meeting the *Frye* test. The Court distinguished this type of evidence, which relies on some scientific principle or test which implies an infallibility not found in pure opinion testimony, from pure opinion testimony which relies solely on the expert's personal experience or training.

The Court stated:

> The jury will naturally assume that the scientific principles underlying the expert's opinion are valid. Accordingly, this type of testimony must meet

> the Frye test, designed to ensure that the jury will not be mislead by experimental scientific methods which may ultimately prove to be unsound.

Id. at 828. The Court held, after reviewing relevant academic literature and case law, that sexual offender profile evidence is not generally accepted in the scientific community and does not meet the *Frye* test for admissibility. The Court cited with approval Judge Ervin's concurring and dissenting opinion in the decision of the First District Court of Appeal.

Similarly, in *Hadden v. State*, 690 So.2d 573 (Fla. 1997), the Florida Supreme Court considered the issue of the admissibility of child sexual abuse accommodation syndrome. The Court held that this syndrome has not been proven by a preponderance of scientific evidence to be generally accepted by a majority of experts in psychology, and therefore does not meet the *Frye* test. The Court again cited with approval Judge Ervin's concurring and dissenting opinion in *Flanagan, supra*, and his dissenting opinion in the First District Court of Appeal decision in *Hadden v. State*, 670 So.2d 77 (Fla. 1st DCA 1996).

In *Hadden*, the Supreme Court held that profile evidence and syndrome evidence suffer from the same infirmity-they have not reached the level of general acceptance in the scientific community. The Court stated:

> We differentiate pure opinion testimony based upon clinical experience from profile and syndrome evidence because profile and syndrome evidence rely on conclusions based upon studies and tests. Further we find that profile or syndrome evidence is not made admissible by combining such evidence with pure opinion evidence because such a combination is not pure opinion evidence based solely upon the expert's clinical experience.(emphasis added)

Id. at 580.

In *Irving v. State*, 705 So.2d 1021 (Fla. 1st DCA 1998), the First District Court of Appeal reaffirmed the principle that expert testimony that a child victim exhibited symptoms consistent with a child who has been sexually abused was inadmissible as not meeting the Frye test. The testifying expert never used the terms "profile" or "syndrome" and instead relied upon post-traumatic stress disorder and related diagnostic criteria. The Court held if the expert bases his opinion on matters other than the expert's experiences, it is subject to the *Frye* test.

Finally, the Florida Supreme Court recently reaffirmed *Flanagan* and *Hadden* in *Williamson v. State*, 994 So.2d 1000 Fla. 2008). *Williamson* held that a *Frye* hearing should have been conducted prior to the admission of expert testimony regarding a state witness displaying "a pattern of someone who has been terrorized." The Court specifically rejected the State's theory that this constituted "pure opinion" testimony and reaffirmed the principle first announced in *Flanagan, supra*, that "in order to introduce expert testimony deduced from a scientific principle or discovery, the principle or discovery 'must be sufficiently established to have gained general acceptance in the particular field in which it belongs'." *Id.* at 1010.

In the case at hand, Dr. Vass's testimony is premised upon a methodology which is not generally accepted by any other expert. Said methodology and the underlying theory behind the methodology has not gained general acceptance in the particular field in which it belongs.

The National Academy of Sciences report on:
"Strengthening Forensic Science in the United States a path forward."

On November 22, 2005, the Science, State, Justice, Commerce and Related Agencies Appropriations Act of 2006 became law. Under the terms of the Statute, Congress authorized "the National Academy of Sciences to conduct a study on forensic science." (NAS report summary pg.1) In 2006, the committee was established which consisted of members from the forensic science community, members of the legal community and a diverse group of scientists. On February 19th 2009, the National Academy of Sciences published their findings and identified numerous areas needed for improvement one of those areas specifically addressed the Courts by stating "the courts continue to rely on forensic evidence without fully understanding and addressing the limitations of different forensic science disciplines. This profound conjunction of law and science, especially in the context of law enforcement, underscores the need for improvement in the forensic science community." The report concludes that "every effort must be made to limit the risk of having the reliability of certain forensic science methodologies judicially certified before the techniques have been properly studied and their accuracy verified." (NAS report pg.85-86)

The National Academy of Sciences recognizes that the laws "greatest dilemma is 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." *Id. at 87.* It recommends that courts look at two questions when admitting forensic evidence (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.

On October 14th 2008, State Attorney of the Ninth Judicial Circuit Lawson Lamar held a press conference and boasted that his office was the first to use DNA evidence in a criminal case, and that his office would again rely on "cutting edge science" to convict Casey Anthony. While this attempt to poison the jury pool did not go unnoticed to the defense, his comparison of Dr. Vass' "sniffer science" to DNA is hardly an accurate one.

DNA evidence unlike many other forensic sciences was developed in the scientific community first, before it was applied to the courts. Gregor Mendel the "Father of Genetics" performed an experiment in 1857 that led to increased interest in the study of genetics. Perhaps Dr Vass' work in 150 years will lead to a title of the "father of chemical decomposition." Prior to that we must remember that DNA technology was not developed overnight from Friedrich Miescher's published method for separating cell nuclei from cytoplasm in 1871 to James Watson and Franci Crik development of the double helix model in 1953 which eventually awarded them the Nobel Prize for Physiology or Medicine for their elucidation of the structure of DNA. Years and years of research from scientists in the United States and Europe followed by multiple well-funded studies (The National Institutes of Health (NIH) and other respected institutions funded and conducted extensive basic research, followed by applied research. Serious studies on DNA analysis preceded the establishment and implementation of "individualization" criteria. *Id at 101.*) led us to the day in 1986 when Sir Alec Jefferies used DNA in a criminal case in Leicestershire, England.[2] Even Orange County prosecutor Tim Berry argued in the Andrews case that the test has been used successfully

---

[2] Joseph Wambaugh (1989) *The Blooding.* New York: Bantam Books.

for 10 years in non-criminal cases, such as paternity suits,[3] the argument making the comparison of this pseudo science to that of DNA would be laughable if it where not such a serious matter.

Chapter 4 of the NAS report addresses the topic of validation of new methods (exhibit A), this can provide courts with a guide as it relates to the arm of science and how it's relationship to the law should be applied.

Defendant specifically requests this Court hold an evidentiary "Frye" hearing on this issue. Thereupon exclude the testimony relating to any air, carpet samples or paper towels tested by Oakridge laboratories in this case.

Wherefore, Defendant requests this Honorable Court exclude the above-referenced testimony or evidence from any trial of this cause.

Respectfully submitted,

ATTORNEYS FOR MISS ANTHONY

JOSE A. BAEZ, ESQUIRE
Florida Bar No.: 013232
The Baez Law Firm
522 Simpson Road
Kissimmee, FL 34744
J. CHENEY MASON, ESQUIRE
Florida Bar No.: 131982
390 North Orange Avenue
Suite 2100
Orlando, Florida 32801-1967
ANN FINNELL, ESQUIRE
Florida Bar No. 0270040
FINNELL, MCGUINNESS,

---

[3] http://www.nytimes.com/1988/02/06/us/rapist-convicted-on-dna-match.html

NEZAMI & ANDUX P.A.
233 E. Bay St. Ste. 601
Jacksonville, Fl. 32202

This description of how science creates new theories illustrates key elements of good scientific practice: precision when defining terms, processes, context, results, and limitations; openness to new ideas, including criticism and refutation; and protections against bias and overstatement (going beyond the facts). Although these elements have been discussed here in the context of creating new methods and knowledge, the same principles hold when applying known processes or knowledge. 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.

Validation of New Methods

One particular task of science is the validation of new methods to determine their reliability under different conditions and their limitations. Such studies begin with a clear hypothesis (e.g., "new method X can reliably associate biological evidence with its source"). An unbiased experiment is designed to provide useful data about the hypothesis. Those data—measurements collected through methodical prescribed observations under well-specified and controlled conditions—are then analyzed to support or refute the hypothesis. The thresholds for supporting or refuting the hypothesis are clearly articulated before the experiment is run. The most important outcomes from such a validation study are (1) information about whether or not the method can discriminate the hypothesis from an alternative, and (2) assessments of the sources of errors and their consequences on the decisions returned by the method. These two outcomes combine to provide precision and clarity about what is meant by "reliably associate."

For a method that has not been subjected to previous extensive study, a researcher might design a broad experiment to assist in gaining knowledge about its performance under a range of conditions. Those data are then analyzed for any underlying patterns that may be useful in planning or interpreting tests that use the new method. In other situations, a process already has been formulated from existing experimental data, knowledge, and theory (e.g., "biological markers A, B, and C can be used in DNA forensic investigations to pair evidence with suspect").

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

The International Organization for Standardization (ISO) and the International Electrotechnical Commission (IEC) developed a joint document,

EXHIBIT 1

"General requirements for the competence of testing and calibration laboratories" (commonly referred to as "ISO 17025"), which includes a well-established list of techniques that can be used, alone or in combination, to validate a method:

- calibration using reference standards or reference materials;
- comparison of results achieved with other methods;
- interlaboratory comparisons;
- systematic assessment of the factors influencing the result; and
- assessment of the uncertainty of the results based on scientific understanding of the theoretical principles of the method and practical experience.[1]

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. Methods that are specified in more detail (such as DNA analysis, where particular genetic loci are to be compared) will have greater credibility and also are more amenable to systematic improvement than those that rely more heavily on the judgments of the investigator.

The validation of results over time increases confidence. Moreover, the scientific culture encourages continued questioning and improvement. Thus, the relevant scientific community continues to check that established results still hold under new conditions and that they continue to hold in the face of new knowledge. The involvement of graduate student researchers in scientific research contributes greatly to this diligence, because part of their education is to read carefully and to question so-called established methods. This culture leads to continued reexamination of past research and hence increased knowledge.

In the case of DNA analysis, studies have evaluated the precision, reliability, and uncertainties of the methods. This knowledge has been used to define standard procedures that, when followed, lead to reliable evidence. For example, below is a brief sample of the specifications required by the Federal Bureau of Investigation's (FBI's) Quality Assurance Standards for

---

[1] Quoted from Section 5.4.5 2 (Note 2) of ISO/IEC 17025, "General requirements for the competence of testing and calibration laboratories" (2nd ed., May 15, 2005).

Forensic DNA Testing Laboratories[2] in order to ensure reliable DNA forensic analysis:

- Testing laboratories must have a standard operating protocol for each analytical technique used, specifying reagents, sample preparation, extraction, equipment, and controls that are standard for DNA analysis and data interpretation.
- The laboratory shall monitor the analytical procedures using appropriate controls and standards, including quantitation standards that estimate the amount of human nuclear DNA recovered by extraction, positive and negative amplification controls, and reagent blanks.
- The laboratory shall check its DNA procedures annually or whenever substantial changes are made to the protocol(s) against an appropriate and available NIST standard reference material or standard traceable to a NIST standard.
- The laboratory shall have and follow written general guidelines for the interpretation of data.
- The laboratory shall verify that all control results are within established tolerance limits.
- Where appropriate, visual matches shall be supported by a numerical match criterion.
- For a given population(s) and/or hypothesis of relatedness, the statistical interpretation shall be made following the recommendations 4.1, 4.2, or 4.3 as deemed applicable of the National Research Council report entitled *The Evaluation of Forensic DNA Evidence* (1996) and/or a court-directed method. These calculations shall be derived from a documented population database appropriate for the calculation.[3]

This level of specificity is consistent with the spirit of the guidelines presented in ISO 17025. The second edition (May 15, 2005) of those guidelines includes the following minimum set of information for properly specifying the process of any new analytical method:

(a) appropriate identification;
(b) scope;
(c) description of the type of item to be tested or calibrated;

---

[2] DNA Advisory Board. 2000. *Forensic Science Communications* 2(3). Available at www.bioforensics.com/conference04/TWGDAM/Quality_Assurance_Standards_2.pdf.

[3] Paraphrased from Section 9 of the FBI's Quality Assurance Standards for Forensic DNA Testing Laboratories.

(d) parameters or quantities and ranges to be determined;
(e) apparatus and equipment, including technical performance requirements;
(f) reference standards and reference materials required;
(g) environmental conditions required and any stabilization period needed;
(h) description of the procedure, including
  - affixing of identification marks, handling, transporting, storing and preparation of items;
  - checks to be made before the work is started;
  - checks that the equipment is working properly and, where required, calibration and adjustment of the equipment before each use;
  - the method of recording the observations and results;
  - any safety measures to be observed;
(i) criteria and/or requirements for approval/rejection;
(j) data to be recorded and method of analysis and presentation;
(k) the uncertainty or the procedure for estimating uncertainty.[4]

Uncertainty and Error

Scientific data and processes are subject to a variety of sources of error. For example, laboratory results and data from questionnaires are subject to measurement error, and interpretations of evidence by human observers are subject to potential biases. A key task for the scientific investigator designing and conducting a scientific study, as well as 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. Numerical data reported in a scientific paper include not just a single value (point estimate) but also a range of plausible values (e.g., a confidence interval, or interval of uncertainty).

Measurement Error

As with all other scientific investigations, laboratory analyses conducted by forensic scientists are subject to measurement error. Such error reflects the intrinsic strengths and limitations of the particular scientific technique. For example, methods for measuring the level of blood alcohol in an individual or methods for measuring the heroin content of a sample

---

[4] Quoted from Section 5.4.4 of ISO/IEC 17025, "General requirements for the competence of testing and calibration laboratories" (2nd ed., May 15, 2005).

can do so only within a confidence interval of possible values. In addition to the inherent limitations of the measurement technique, a range of other factors may also be present and can affect the accuracy of laboratory analyses. Such factors may include deficiencies in the reference materials used in the analysis, equipment errors, environmental conditions that lie outside the range within which the method was validated, sample mix-ups and contamination, transcriptional errors, and more.

Consider, for example, a case in which an instrument (e.g., a breathalyzer such as Intoxilyzer) is used to measure the blood-alcohol level of an individual three times, and the three measurements are 0.08 percent, 0.09 percent, and 0.10 percent. The variability in the three measurements may arise from the internal components of the instrument, the different times and ways in which the measurements were taken, or a variety of other factors. These measured results need to be reported, along with a confidence interval that has a high probability of containing the true blood-alcohol level (e.g., the mean plus or minus two standard deviations). For this illustration, the average is 0.09 percent and the standard deviation is 0.01 percent; therefore, a two-standard-deviation confidence interval (0.07 percent, 0.11 percent) has a high probability of containing the person's true blood-alcohol level. (Statistical models dictate the methods for generating such intervals in other circumstances so that they have a high probability of containing the true result.) The situation for assessing heroin content from a sample of white powder is similar, although the quantification and limits are not as broadly standardized. The combination of gas chromatography and mass spectrometry (GC/MS) is used extensively in identifying controlled substances. Those analyses tend to be more qualitative (e.g., identifying peaks on a spectrum that appear at frequencies consistent with the controlled substance and which stand out above the background "noise"), although quantification is possible.

Error Rates

Analyses in the forensic science disciplines are conducted to provide information for a variety of purposes in the criminal justice process. However, most of these analyses aim to address two broad types of questions: (1) can a particular piece of evidence be associated with a particular class of sources? and (2) Can a particular piece of evidence be associated with one particular source? The first type of question leads to "classification" conclusions. An example of such a question would be whether a particular hair specimen shares physical characteristics common to a particular ethnic group. An affirmative answer to a classification question indicates only that the item belongs to a particular *class* of similar items. Another example might be whether a paint mark left at a crime scene is consistent (according to some collection of relevant measurements) with a particular paint sample in a database, from which one can infer the class of vehicle (e.g., model(s) and production year(s)) that could have left the mark. The second type of question leads to "individualization" conclusions—for example, does a particular DNA sample belong to individual X?

Although the questions addressed by forensic analyses are not always binary (yes/no) or as crisply stated as in the previous paragraph, the paradigm of yes/no conclusions is useful for describing and quantifying the accuracy with which forensic science disciplines can provide answers.[5] In such situations, results from analyses for which the truth is known can be classified in a two-way table as follows:

	Analysis Results	
Truth	yes	no
yes	a (true positives)	b (false negatives)
no	c (false positives)	d (true negatives)

The conceptual framework and terminology for evaluating the accuracy of forensic analyses is illustrated using a *hypothetical* example from microscopic analysis of head hair. In this situation, multiple features, both qualitative and quantitative, on each sample of hair are assessed. Qualitative features include color (e.g., blonde, brown, red), coloring (natural or treated), form (straight, wavy, curved, kinked), texture (smooth, medium, coarse). Quantitative features include length and diameter. Undoubtedly, these features will vary from hair to hair, even from the same individual, but features that vary *less* for the *same* individual (i.e., within-individual variability) and *more* for *different* individuals (i.e., between-individual variability) are needed for purposes of class identification and discrimination. These features may also be combined in some fashion to result in some overall score, or set of scores, for each sample, and these scores are then compared with those from the target sample. In the final analysis, however, a binary conclusion is often required. For example, "Did this hair come from the head of a Caucasian person?"

As in the case of all analyses leading to classification conclusions (e.g., diagnostic tests in medicine), the microscopic hair analysis process must be subjected to performance and validation studies in which appropriate error rates can be defined and estimated. Consider a hypothetical study in

---

[5] More complete discussion of the questions addressed by forensic science may be found in references such as K. Inman and N. Rudin. 2002. The origin of evidence. *Forensic Science International* 126:11-16; and R. Cook, I.W. Evett, G. Jackson, P.J. Jones, and J.A. Lambert. 1998. A hierarchy of propositions: Deciding which level to address in casework. *Science and Justice* 38:231-239.

which 100 samples (each with multiple hairs) are taken from the heads of 100 individuals from class C, and another 100 samples are taken from the heads of individuals not in class C. The analyst is asked to determine, for each of the 200 samples, whether it does or does not come from a person in class C, and the true answer is known. The validation study returns the following results:

Hypothetical Hair Analysis Validation Study

	Analysis of Hair Samples Indicates:		
	Class C	Not Class C	Row Total
Sample is from Class C Persons	95 True Positive (correct determination)	5 False Negative	100
Sample is not from Class C Persons	2 False Positive	98 True Negative (correct determination)	100
Column Total	97	103	Overall total 200

The accuracy of a test (here, microscopic hair analysis) can be assessed in different ways. Borrowing terminology from the evaluation of medical diagnostic tests, four characterizations and their associated measures are given below. Each one is useful in its own way: the first two emphasize the ability to *detect* an association; the last two emphasize the ability to *predict* an association:[6]

- Among samples from persons in Class C, the fraction that is correctly identified by the test is called the "sensitivity" or the "true positive rate" (TPR) of the test. In this table, the sensitivity would be estimated as [95/(95+5)] × 100=95 percent.
- Among samples from persons not in Class C, the fraction that is correctly identified by the test is called the "specificity" or the "true

---

[6] See, e.g., X-H. Zhou, N. Obuchowski, and D. McClish. 2002. *Statistical Methods in Diagnostic Medicine.* Hoboken, NJ. Wiley & Sons, for a general account of methods for diagnostic tests. A series of NAS/NRC reports have applied such methods to the examination of forensic disciplines. See, e.g., NRC. Committee to Review the Scientific Evidence on the Polygraph. 2003. *The Polygraph and Lie Detection.* Washington, DC: The National Academies Press; NRC. 2004. *Forensic Analysis: Weighing Bullet Lead Evidence.* Washington, DC: The National Academies Press; NAS. 2005. *The Sackler Colloquium on Forensic Science: The Nexus of Science and the Law*, November 16-18, 2005.

negative rate" (TNR) of the test. In this table, the specificity would be estimated as [98/(2+98)] × 100=98 percent.

- Among samples classified by the test as coming from persons in Class C, the fraction that actually turns out to be from Class C is called the "positive predictive value (PPV)" of the test. In this table, the PPV would be estimated as [95/(95+ 2)] × 100=98 percent.
- Among samples classified by the test as coming from persons not in Class C, the fraction that actually turns out to not be persons from Class C is called the "negative predictive value (NPV)" of the test. In this table, the NPV would be estimated as [98/(5+98)] × 100=95 percent.

The above four measures emphasize the ability of the analysis to make correct determinations.[7] "Error rates" are defined as proportions of cases in which the analysis led to a false conclusion. For example, the complement of sensitivity (100 percent minus the sensitivity) is the percent of false negative cases in which the sample was from class C but the analysis reached the opposite conclusion. In the above table, this would be estimated as 5 percent. Similarly, the complement of specificity (100 percent minus the specificity) is the percent of false positive cases in which the sample was not from class C but the analysis concluded that it was. In the above table this would be estimated as 2 percent. A global error rate could be defined as the percent of incorrectly identified cases among all those analyzed. In the above table this would be estimated as [(5+2)/200] × 100=3.5 percent.

Importantly, whether the test answer is correct or not depends on which question is being addressed by the test. In this hair comparison example, the purpose is to determine whether the hair came from the head of an individual from class C. Thus, the analysis should be evaluated on the accuracy of the classification. In this example, if the analysis indicated "Class C" but the hair actually came from a "non-Class C" individual, then the analysis returned an incorrect classification. This accuracy evaluation does not apply to other tasks that are beyond the goal of the particular analysis, such as pinpointing the individual from whom the specimen was obtained. In the paint example about paint marks left by a vehicle, if the question is whether a vehicle under investigation was a model A made by manufacturer B in 2000, then a correct answer is limited to only the model, manufacturer, and year.

---

[7] Each estimate (of sensitivity, specificity, PPV, NPV) is associated with an interval that has a high probability of containing the true sensitivity, specificity, PPV, NPV. The larger the study, the more precise the estimate (i.e., the narrower the interval of uncertainty about the estimate).

Although only illustrations, these examples serve to demonstrate the importance of:

- the careful and precise characterization of the scientific procedure, so that others can replicate and validate it;
- the identification of as many sources of error as possible that can affect both the accuracy and precision of a measurement;
- the quantification of measurements (e.g., in the example of GC/MS analysis of possible heroin, reporting peak area, as well as appropriate calibration data, including the response area for a known amount of analyte standard, rather than merely "peak is present/absent");
- the reporting of a measurement with an interval that has a high probability of containing the true value;
- the precise definition of the question addressed by the method (e.g., classification versus individualization), and the recognition of its limitations; and
- the conducting of validation studies of the performance of a forensic procedure to assess the percentages of false positives and false negatives.

Clearly, better understanding of the measuring equipment and the measurement process leads to more improvements to every process and ultimately to fewer false positive and false negative results. Most importantly, as stated above, whether the test answer is correct or not depends on the question the test is being used to address. In the case of microscopic hair analysis, the validation study may confirm its value in identifying *class characteristics* of an individual, but not in identifying the specific person.

It is also important to note that errors 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.[8]

---

[8] C. Aitken and F. Taroni. 2004. *Statistics and the Evaluation of Evidence for Forensic Scientists*. Chichester, UK: John Wiley & Sons.

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. The estimation of such error rates requires rigorously developed and conducted scientific studies. Additional factors may play a role in analyses involving human interpretation, such as the experience, training, and inherent ability of the interpreter, the protocol for conducting the interpretation, and biases from a variety of sources, as discussed in the next section. 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.

Sources of Bias

Human judgment is subject to many different types of bias, because we unconsciously pick up cues from our environment and factor them in an unstated way into our mental analyses. Those mental analyses might also be affected by unwarranted assumptions and a degree of overconfidence that we do not even recognize in ourselves. Such cognitive biases are not the result of character flaws; instead, they are common features of decisionmaking, and they cannot be willed away.[9] A familiar example is how the common desire to please others (or avoid conflict) can skew one's judgment if co-workers or supervisors suggest that they are hoping for, or have reached, a particular outcome. Science takes great pains to avoid biases by using strict protocols to minimize their effects. The 1996 National Academies DNA report, for example, notes, "[l]aboratory procedures should be designed with safeguards to detect bias and to identify cases of true ambiguity. Potential ambiguities should be documented."[10]

A somewhat obvious cognitive bias that may arise in forensic science is a willingness to ignore base rate information in assessing the probative value of information. For example, suppose carpet fibers from a crime scene are found to match carpet fibers found in a suspect's home. The probative value of this information depends on the rate at which such fibers are found in homes in addition to that of the suspect. If the carpet fibers are extremely common, the presence of matching fibers in the suspect's home will be of little probative value.[11]

A common cognitive bias is the tendency for conclusions to be affected by how a question is framed or how data are presented. In a police line-up,

---

[9] See, e.g., M.J. Saks, D.M. Risinger, R. Rosenthal, and W.C. Thompson. 2003. Context effects in forensic science: A review and application of the science of science to crime laboratory practice in the United States. *Science and Justice* 43(2):77-90.

[10] NRC. 1996. *The Evaluation of Forensic DNA Evidence.* Washington, DC: National Academy Press.

[11] C. Guthrie, J.J. Rachlinski, and A.J. Wistrich. 2001. Inside the judicial mind. *Cornell Law Review* 86:777-830.

for instance, an eyewitness who is presented with a pool of faces in one batch might assume that the suspect is among them, which may not be correct. If the mug shots are presented together at one time and the witness is asked to identify the suspect, the witness may choose the photograph that is most similar to the perpetrator, even if the perpetrator's picture is not among those presented. Similarly, if the photographs are presented sequentially and the witness knows that only a limited number will be presented, the eyewitness might tend to "identify" one of the last photographs under the assumption that the suspect must be in that batch. (This is also driven by the common bias toward reaching closure.) A series of studies has shown that judges can be subject to errors in judgment resulting from similar cognitive biases.[12] Forensic scientists also can be affected by this cognitive bias if, for example, they are asked to compare two particular hairs, shoeprints, fingerprints—one from the crime scene and one from a suspect—rather than comparing the crime scene exemplar with a pool of counterparts.

Another potential bias is illustrated by the erroneous fingerprint identification of Brandon Mayfield as someone involved with the Madrid train bombing in 2004. The FBI investigation determined that once the fingerprint examiner had declared a match, both he and other examiners who were aware of this finding were influenced by the urgency of the investigation to affirm repeatedly this erroneous decision.[13]

Recent research provided additional evidence of this sort of bias through an experiment in which experienced fingerprint examiners were asked to analyze fingerprints that, unknown to them, they had analyzed previously in their careers. For half the examinations, contextual biasing was introduced. For example, the instructions accompanying the latent prints included information such as the "suspect confessed to the crime" or the "suspect was in police custody at the time of the crime." In 6 of the 24 examinations that included contextual manipulation, the examiners reached conclusions that were consistent with the biasing information and different from the results they had reached when examining the same prints in their daily work.[14]

Other cognitive biases may be traced to common imperfections in our reasoning ability. One commonly recognized bias is the tendency to avoid cognitive dissonance, such as persuading oneself through rational argument that a purchase was a good value once the transaction is complete. A scientist encounters this unconscious bias if he/she becomes too wedded to a preliminary conclusion, so that it becomes difficult to accept new infor-

---

[12] Ibid.

[13] R.B. Stacey. 2004. A report on the erroneous fingerprint individualization in the Madrid train bombing case. *Journal of Forensic Identification* 54:707.

[14] I.E. Dror and D. Charlton. 2006. Why experts make errors. *Journal of Forensic Identification* 56(4):600-616.

mation fairly and unduly difficult to conclude that the initial hypotheses were wrong. This is often manifested by what is known as "anchoring," the well-known tendency to rely too heavily on one piece of information when making decisions. Often, the piece of information that is weighted disproportionately is one of the very first ones encountered. One tends to seek closure and to view the initial part of an investigation as a "sunk cost" that would be wasted if overturned.

Another common cognitive bias is the tendency to see patterns that do not actually exist. This bias is related to our tendency to underestimate the amount of complexity that can really exist in nature. Both tendencies can lead one to formulate overly simple models of reality and thus to read too much significance into coincidences and surprises. More generally, human intuition is not a good substitute for careful reasoning when probabilities are concerned. As an example, consider a problem commonly posed in beginning statistics classes: How many people must be in a room before there is a 50 percent probability that at least two will share a common birthday? Intuition might suggest a large number, perhaps over 100, but the actual answer is 23. This is not difficult to prove through careful logic, but intuition is likely to be misleading.

All of these sources of bias are well known in science, and a large amount of effort has been devoted to understanding and mitigating them. The goal is to make scientific investigations as objective as possible so the results do not depend on the investigator. Certain fields of science (most notably, biopharmaceutical clinical trials of treatment protocols and drugs) have developed practices such as double-blind tests and independent (blind) verification to minimize the impact of biases. Additionally, science seeks to publish its discoveries, findings, and conclusions so that they are subjected to independent peer review; this enables others to study biases that may exist in the investigative method or attempt to replicate unexpected results. Avoiding, or compensating for, a bias is an important task. Even fields with well-established protocols to minimize the effects of bias can still bear improvement. For example, a recent working paper[15] has raised questions about the way cognitive dissonance has been studied since 1956. Although these results must be considered preliminary because the paper has yet to be published, they do demonstrate that continual vigilance is needed. Research has been sparse on the important topic of cognitive bias in forensic science—both regarding their effects and methods for minimizing them.[16]

---

[15] M.K. Chen. 2008. *Rationalization and Cognitive Dissonance: Do Choices Affect or Reflect Preferences?* Available at www.som.yale.edu/Faculty/keith.chen/papers/CogDisPaper.pdf.

[16] See, e.g., I.E. Dror, D. Charlton, and A.E. Peron. 2006. Contextual information renders experts vulnerable to making erroneous identifications. *Forensic Science International* 156:74-78; I.E. Dror, A. Peron, S. Hind, and D. Charlton. 2005. When emotions get the better of us:

The Self-Correcting Nature of Science

The methods and culture of scientific research enable it to be a self-correcting enterprise. Because researchers are, by definition, creating new understanding, they must be as cautious as possible before asserting a new "truth." Also, because researchers are working at a frontier, few others may have the knowledge to catch and correct any errors they make. Thus, science has had to develop means of revisiting provisional results and revealing errors before they are widely used. The processes of peer review, publication, collegial interactions (e.g., sharing at conferences), and the involvement of graduate students (who are expected to question as they learn) all support this need. Science is characterized also by a culture that encourages and rewards critical questioning of past results and of colleagues. Most technologies benefit from a solid research foundation in academia and ample opportunity for peer-to-peer stimulation and critical assessment, review and critique through conferences, seminars, publishing, and more. These elements provide a rich set of paths through which new ideas and skepticism can travel and opportunities for scientists to step away from their day-to-day work and take a longer-term view. The scientific culture encourages cautious, precise statements and discourages statements that go beyond established facts; it is acceptable for colleagues to challenge one another, even if the challenger is more junior. The forensic science disciplines will profit enormously by full adoption of this scientific culture.

CONCLUSION

The way in which science is conducted is distinct from, and complementary to, other modes by which humans investigate and create. The methods of science have a long history of successfully building useful and trustworthy knowledge and filling gaps while also correcting past errors. The premium that science places on precision, objectivity, critical thinking, careful observation and practice, repeatability, uncertainty management, and peer review enables the reliable collection, measurement, and interpretation of clues in order to produce knowledge.

The effects of contextual top-down processing on matching fingerprints. *Journal of Applied Cognitive Psychology* 19:799-809; and B. Schiffer and C. Champod. 2007. The potential (negative) influence of observational biases at the analysis stage of fingerprint individualization. *Forensic Science International* 167:116-120.