UNITED STATES OF AMERICA
FOR THE DISTRICT OF MARYLAND
SOUTHERN DIVISION

UNITED STATES OF AMERICA,	:	Criminal Action No.:
vs.	:	PWG 17-242
JOVON MEDLEY,	:	Greenbelt, Maryland
Defendant.	:	Tuesday, April 24, 2018

_____/

TRANSCRIPT OF EXCERPT OF MOTION HEARING PROCEEDINGS
BEFORE THE HONORABLE PAUL W. GRIMM
UNITED STATES DISTRICT JUDGE

APPEARANCES:

FOR THE GOVERNMENT: CHRISTIAN J. NAUVEL, ESQUIRE
LINDSAY EYLER KAPLAN, ESQUIRE
U.S. Attorney's Office
6500 Cherrywood Lane
Suite 200
Greenbelt, MD 20770
(301) 344-0113

FOR THE DEFENDANT; MICHAEL S. OPPENHEIMER, ESQUIRE
CHRISTIAN BRYAN LASSITER, ESQUIRE
Office of the Federal Public Defender
100 S Charles Street
Tower II, Ninth Floor
Baltimore, MD 20201
(410) 962-3962

OFFICIAL COURT REPORTER: LINDA C. MARSHALL, (301) 344-3229
COMPUTER-AIDED TRANSCRIPTION OF STENOTYPE NOTES

P-R-O-C-E-E-D-I-N-G-S

EXCERPT

(Whereupon, other proceedings were reported, but are not herein transcribed.)

THE COURT: All right. Are we ready to proceed?

MR. OPPENHEIMER: Yes, Your Honor.

MR. NAUVEL: Yes, Your Honor.

THE COURT: All right. The two operative filings in this case are ECF 68, Defense Motion In Limine to Exclude Firearm Identification Evidence or in the Alternative to Limit it, which was filed on March 21st, 2018. And ECF 78, the government's response to this motion, ECF 78 filed on April 9th, 2018.

Each of the motions contained a significant number of exhibits, the defense exhibits were numbered A through N, and include ballistic imaging report excerpts, excerpts from the PCAST, which is President's Council of Advisors for Science and Technology, excerpts, PCAST addendum, the National Research Committee's forensic report, the disclosure of Mr. McVeigh in this case, the affidavit of Mr. Nixon, who is the defense expert, a article from "Jurimetrics," J-U-R-I-M-E-T-R-I-C-S, 51 "Jurimetrics Journal," summer 2010; the AFTE, which is the Association of Firearms and Tool Mark Examiners, training module; the AFTE, theory of identification; a OSAC research needs assessment form, and -- hold on one second.

(Pause.)

THE COURT: Department of Justice memorandum from Attorney General Lynch dated September 2016 regarding forensic evidence disclosures, firearm and tool mark discovery requests submitted by the defense, the government's discovery production letter, and the Department of Justice Code of Professional Responsibility for the Practice of Forensic Science.

In response, the exhibits produced by the government were the Prince George's County police investigator's notes from Mr. McVeigh, witness statements, the crime scene report, various documents associated with the NIBIN, N-I-B-I-N, analysis, which was testified about yesterday. Firearms analysis, the response of the AFTE to the National Academy of Sciences 2008 report addressing feasibility, accuracy and technical capability of national ballistic database. The response to 25 foundational firearm and tool mark examination questions, the 2016 PCAST report, the OSAC response to the PCAST report.

Give me one second. Yes, finally, OSAC, Organization of Scientific Area Committees, the tool marks, the firemarks -- Firearms and Tool Mark Subcommittee dated 14 December 2016, which was submitted to PCAST in response to its December 2016 report. And then various documents with regard to the discovery produced in this case at the hearing. Additional documents were introduced into evidence and received as well.

I want to start with the general evidentiary issue

that we have to address. From there I want to deal with past judicial opinions that have focused on firearms evidence. I then want to focus on the more recent analysis of firearms, tool mark evidence inaugurated by the 2015 PCAST report and subsequently. Then I want to address the specific expert report findings in this particular case and leave you with my ruling.

The challenge as presented by the defense comes from Federal Rule of Evidence 702 that deals with expert testimony. That rule, which was last codified in December of 2000, says that, in essence, when if scientific technical or specialized information will assist the fact-finder in understanding the evidence or making a determination of an issue that is left to the fact-finder, then an expert qualified by virtue of knowledge, training, experience, background, education or skill may testify in the form of an opinion or otherwise. That was the basic formula up until 2000 that had been in existence since the mid-1970s when the Rules of Evidence were first codified.

But in response to the trilogy of cases from the mid-1990s to the end of the 1990s in *Daubert versus Merrell Dow*, the *Kumho Tire*, I think it's *Kumho Tire versus Carmichael*, and the *Joiner* case, the Advisory Committee for the Rules of Evidence recommended and the Supreme Court and Congress approved changes to Rule 702 that added the new language after the introductory clause I have just paraphrased. That new language required that the expert could testify in a form of an opinion

or otherwise if properly qualified, and if it would be helpful to the jury. If, one, the expert's opinion was based upon sufficient facts or data.

Two, the methods and principles relied upon were reliable.

And three, the methods and principles used would reliably apply to the facts of the particular case.

Those final three qualifications were the product of the Supreme Court's discussions in the *Daubert*, *Joiner* and *Kumho Tire*, *K-U-M-H-O, Tire* cases.

The new area of evidence law that was inaugurated by *Daubert* in the mid-1990's thrust upon federal judges a role that many feel they're ill equipped for. And that is, judges who are generalists are required to make specific assessments regarding the validity and reliability of underlying scientific and technical processes that come into court very frequently.

Daubert was an interesting case because it was a method/epidemiological case involving the birth control drug Bendectin -- not birth control, but the antinausea drug for pregnant women, Bendectin, and whether it caused birth defects. And the plaintiff who had very qualified experts had introduced in response to the defense Motion for Summary Judgment the affidavits of a number of very qualified experts who expressed the opinion that based upon chemical analyses in lab testing of rodents, and the careful analysis of the underlying data in

various published opinions that themselves did not find the causal connection, that there was a causal connection between Bendectin and birth defects.

The defendants had received that expert disclosure as required by Federal Rule of Civil Procedure 26(a)(2), and when discovery had concluded, they filed a Motion for Summary Judgment with an affidavit of a medical doctor who was an epidemiologist who swore in his affidavit that he had reviewed all 300 plus published articles that dealt with Bendectin, and that there was no epidemiologically recognized causal connection between Bendectin and birth defects. And as a result of that, the District Court excluded the plaintiff's experts and granted summary judgment in favor of the plaintiffs -- or the defendant, excuse me.

It went up on appeal and ultimately the Supreme Court granted cert. And largely the epidemiological affidavit of the doctor had taken the position that the methodology used by the plaintiff's doctor in *Daubert* was flawed because the generally accepted methodology for determining causation of health, adverse health effects in the large population of people came not from clinical medicine, the assessment of individual patients' health and the treatment thereof, but rather epidemiological studies of broad populations throughout a lengthy period of time, and that that methodology used by the plaintiff's experts was not generally accepted as reliable among

the relevant scientific community, which was epidemiologists.

At the Supreme Court, the plaintiffs and their amici shocked all of the observers because they took the position that this general accepted standard which came from *United States versus Frye*, a decision of the 1920s by the District of Columbia Court of Appeals that rejected the proposed testimony of a precursor to a lie detector test.

The plaintiffs in their amici argued that that general accepted standard had no place in Evidence Rule 702, which when the Rules of Evidence were codified in the mid-1970s supplanted that common law approach for determining when novel science was admissible.

And what *Frye* had said was, "In the continuum of science, which marches on for all times, no individual knows," and this is paraphrasing language actually used by the court in that twilight zone of science, and they used the phrase "twilight zone," an idea which has promise crosses the line and becomes generally accepted by the scientific community. But when it does, it's reliable enough to be brought to the attention of juries to be used in deciding cases.

And the Supreme Court was faced with the challenge of whether or not there was any position for the *Frye* case as a standard when Rule 702 was absolutely silent about whether that standard applied.

The Supreme Court's *Daubert* decision started from the

broad propositions that they felt were important, and there were four of them: Relevance, reliability, helpfulness and fit. Evidence is not relevant under Rule 401, has no tendency to prove something is important to the case and shouldn't be admitted.

Reliable evidence is not -- unreliable evidence is not relevant. If methodology does not produce reliable results, it has no tendency to prove something important in the case. Reliability is key.

Helpfulness is key for expert testimonies. We want the jury to be helped. Unreliable expert testimony, methodology that's not reliable is not relevant and, therefore, can't help.

And then there's fit, whether or not the methodology used fits the facts of that particular case. And from these four overlapping areas of evidentiary concern, the Supreme Court moved on to say that there would be a non-exclusive five factor test that federal judges were required to employ under Federal Rule of Evidence 104(a) that says that the judge makes preliminary determinations regarding the admissibility of evidence, the qualification of witnesses and the existence of a privilege. And in doing so is not required to strictly apply the Rules of Evidence except for privilege. That, by the way, is bolstered by Evidence Rule 1101(d)(1), which says that the facts upon which the court makes a 104(a) ruling are not strictly bound by the Rules of Evidence with the exception of

privilege.

And judges, after *Daubert*, were supposed to exercise this gatekeeper function. And to do so there were five non-exclusive tests that were identified.

Number one, has the methodology been tested.

Number two, is there a known error rate.

Number three, has the methodology and the testing of that been subject to peer review evaluation.

And the court put great faith in peer review as it occurs in the scientific community when research findings are published, a board of distinguished peer reviewers evaluates the -- the nature and the design of the experiment ensures that it is appropriate for publication. And then the publication process draws review by other scientists in the field, criticism support, spinoff experiments. And from that process science is supposed to evolve identifying promising reliable methodology and eliminating that which, perhaps, promising but turned out not to have sustainability when subject to subsequent review.

Third -- or fourth was whether or not the methodology had generally been accepted by the relevant community.

And the community here is scientific, technical or specialized. Manifestly not judicial as the Seventh Circuit and the Sixth Circuit in other context, looking at cell tower information, which is our next hearing, have said, the judicial community and the law enforcement community are not the proper

communities for determining whether scientific, technical or specialized information is reliable. That is a community of those who practice that particular science, technology or their technical discipline.

And the mere facts that -- the mere fact that courts have relied upon a certain methodology and admitted it in the past does not grandfather that methodology once the *Daubert* decision and Rule 702 was changed to require judicial independent analysis.

The final *Daubert* factor was whether or not there were standard tests and protocols that governed the methodology and whether they had been complied with. So it's testing, error rate, peer review, general acceptance and compliance with standard testing protocol.

As I mentioned when counsel were capably arguing, the judicial community was gobsmacked by this decision. And judges, as the Chief Justice Rehnquist in his -- in his separate opinion that he wrote in *Daubert* said essentially that he had no -- he stood second to no person in his respect for the federal judiciary, but how in the world federal judges, who are generalists, were supposed to apply falsifiability to a myriad of scientific and technical disciplines in the context of resolving a specific case was a mystery to him.

Notwithstanding that, the mystery continues, and after the *Daubert* case, the *Kumho Tire* case applied the *Daubert*

methodology to nonspecifically scientific disciplines. That involved explosion of a tire, and it was more a matter of engineering and manufacturing than it was science. And then the *Joiner* decision also provided insight.

So the efforts of the judiciary to confine like putting the Jeannie in the bottle, the analysis of *Daubert* to hard science was dispelled after that trilogy of Supreme Court cases. The changes to Rule 702 were adopted, and since that time courts have been required to allow it. And to follow that procedure.

There's not only associated with *Daubert* because *Daubert* was actually intended to liberalize the introduction of scientific, technical and specialized information because previously, the hammerlock of general acceptance meant that there could be very promising and potentially reliable methodology that had not been around long enough to obtain or to attain general acceptance that would be rejected because they weren't generally accepted. And *Daubert* was actually conceived of as a liberalizing analysis to be used.

Experience has shown it's been used primarily just for the opposite, namely it tried to exclude evidence, but regardless of whether it's used to include or exclude evidence, *Daubert* puts a substantial burden on the court to exercise that gatekeeping function. And it is not permissible for courts to ignore that gatekeeping function when faced with challenges that

raise into question the various elements of what *Daubert* requires.

Turning my attention now to the firearms, tool mark evidence. The starting point for my analysis is my report and recommendation, which was adopted by Judge Quarles, when I was a Magistrate Judge, eight years ago in *United States versus Willock*, W-I-L-L-O-C-K, 696 F.Supp. 2d 536, which dealt with firearms, tool marks analysis. I am not going to verbatim go through every bit of that decision, but I'm going to talk about some of the significant concepts in that opinion that continue to have liability today.

In 2010, as for tool mark evidence, particularly being used primarily in criminal cases to identify whether or not bullet fragments, bullets or cartridge casings were fired from a particular firearm that was involved in a crime, that evidence had been around for probably a hundred years and had been widely accepted without any real challenge or analysis by state and federal courts throughout the country and the world.

And, indeed, as evidence of that, when I asked him this question yesterday, Mr. McVeigh testified that in the 80 times in which he had testified as an expert regarding his discipline, only twice, this case and one other case, had there ever been a real challenge that required the court to exercise its function. And that's a perfect reason to explain why the mere fact that courts in the past have accepted certain

technical or specialized information is absolutely not in itself sufficient for courts to continue to admit it. It's just not allowed. It's not what *Daubert* requires, although it is easy sometimes just to say, well, those other courts have done it so will I.

As pointed out in *Willock*, at that particular time, 2010, the National Research Council at the request of the National Academy of Sciences had begun to look at forensic evidence and had published a report that focused on many of the forensic comparative forensic analysis and had reached -- had expressed opinions regarding whether it was reliable.

And in the *Willock* case, I cited to all of the significant federal court decisions that had been decided as of that date to have looked at the admissibility of tool mark evidence. Obviously in 2010, that was ten years after the changes to Rule 702. Some courts had done what I will call rigorous 702 analysis, others had not, had just simply accepted this evidence because it had always been introduced in the past.

The basic foundation of tool mark forensic evidence was discussed in the *Willock* case beginning at page 555 of the opinion. Essentially what it says is a tool mark is a mark that's generated when a hard object known as the tool comes into contact with a relatively softer object.

And that, for example, in the context relevant to this case is when a firearm makes contact with -- when it's made,

when you have a hard tool that gouges or carves or removes metal in order to form the component parts of the firearm; the barrel, the chamber, the breach face, the firing pin, and if you're dealing with a semi-automatic as we are here, the extractor, the ejector. And when the tools that manufacture these manufactured products, the components of the firearm are softer than the tool that makes them by definition and the thought is that microscopic magnification, you can see marks made by the tools on the item that was manufactured.

Similarly, when a semi-automatic handgun is discharged, the theory goes, marks are transferred to the components of the ammunition.

So, for example, when a semi-automatic handgun advances a cartridge which composed of the -- the cartridge casing, the bullet, and inside of it is the explosive material that propels the bullet in the back of which is the firing pin, it comes up into the chamber. When the -- when the slide of the handgun is pulled back and advances it into the chamber and goes forward, the breach face comes into contact with the back of the cartridge. When the trigger is pulled, the firing pin is impacted by the hammer, which can leave an impression in the center of the rim of the back of the cartridge.

The explosion that occurs when the firing pin detonates the -- ignites the explosive, propels the bullet, which goes through the rifling in the barrel, which can go from

left to right or right to left, expels in a spinning fashion the bullet, and the bullet then goes to wherever it lands.

The gas that is expelled by the explosion then forces the ejector back, it then grabs with another piece of the mechanism the rim of the cartridge, flips it out, and advances another one up into the chamber, and the process starts again as quickly as the trigger can be pulled.

And the theory is, is that those forces cause the harder metal of the component parts of the semi-automatic handgun to transmit marks on the bullet made by the lands and grooves of the barrel and the cartridge as it is grabbed and pushed and pulled and ejected.

And those characteristics, the theory goes, can be microscopically looked at, and you can compare cartridges fired or bullets fired from the same handgun over a succession of time.

And the theory is that if tool marks are unique and transmitted uniquely by each gun onto the components of the bullet when it is fired, then you would expect to see the same matching set of marks on the components of a bullet if you examined it under high magnification. That's the theory.

The National Research Council's Committee that was charged in 2008 in this report reported on its examination of ballistic tool mark analysis, and one of the things that is associated with it is sort of three definitions that become key

in the way in which the science and technology is supposed to operate.

First of all, you have class characteristics. Those are family resemblances which will be present in all weapons of the same make and model. So we have the .45 caliber as distinct from the .38 caliber or a .32. We have a revolver which operates differently than a semi-automatic. We heard testimony from Mr. McVeigh yesterday that some rifling goes to -- toward, twists toward the right. Other rifling may twist toward the left. There may be different numbers of spirals in the rifling, the longer the barrel, the more the spirals, and those are class characteristics.

So if, for example, you find a .38 caliber shell at a crime scene, then by definition you've eliminated every larger caliber and smaller caliber handgun from the likely candidates that could have fired that gun. The .45 eliminates a .32 and .38 and a 9-millimeter and those are class characteristics.

Class characteristics that can cause impressions on -- on bullets and casings and cartridge casings including the caliber, type of breach face, type of firing pin, and the breach face can be parallel, arched, smooth or granular or circular. The firing pin can leave an impression that's circular or rectangular or elliptical.

Then we have subclass characteristics. Subclass characteristics are defined, and this is at page 58, by the way,

of the *Willock* case, produce incident to the manufacturer and can arise from a source when changes over time and therefore may be present on a group of guns within certain make or model such as those manufactured at a particular time and place. That's the *Montiero* case, the *Diaz* case and the NRC ballistic imaging report cited at page 558.

So, the notion would be that an example of subclass characteristics would include imperfections that are present on the tool that creates the component parts of a firearm that, for example, are manufactured in a single manufacturing run. So we've got the same kind of product that is being used to make the barrel, for example. We've got the same tools that are used to make the barrel, and we're going to run all these through in one run of 200 barrels.

And the theory is, is that those will have subclass characteristics. So they'll be .45 caliber, they'll be barrel that has a right groove twist. There'll be seven curves in the rifling, but each one having been manufactured by the same tools will bear upon microscopic examination unique marks that were imprinted on them by that particular manufacturing process.

And the theory is, is that if you were to fire ammunition through all of those, then you would expect to see some marks on all of them that were similar because they were manufactured from products all created by the same tools in the same individual manufacturing process. That's the theory

anyway. And that's discussed by Mr. Nichols in his 52 "Journal of Forensic Science" article at page 587, 2007.

Finally, we have individual characteristics, which are defined as random imperfections produced either during the manufacture or by accidental damage which are unique to the object and distinguish it from all others.

The -- and this, by the way, is sort of the key, an analytical issue of tool mark analysis in firearms cases. And that is whether or not, when you look at the marks on a shell casing found at a crime scene, whether they were, they bear the identity of individual characteristics unique to them based upon the unique marks on the handgun that fired them or whether they are -- bear the characteristics of marks that they share as subclass characteristics from a number of guns manufactured by the same tool. And that is the whole underlying issue as to whether or not the firearms examiner can state that the marks on the bullet or the cartridge are unique and individual to that particular firearm to the exclusion of other firearms.

Now, how does a firearms examiner get from the examination of a set of cartridges and bullets found at a crime scene to a decision as to whether or not they were fired by one particular individual firearm that has been proven to be connected to the defendant?

They do so through a process that allows them to reach the decision that the marks on the cartridges or the bullet

fragments found at the crime scene are in sufficient agreement with the marks that are found on cartridges and bullet components that have been known to have been produced from the handgun at issue. And page 560 of *Willock*, the definition of sufficient agreement that was used in the *Montiero* case that was quoting from the AFTE, which is the Association of Firearms and Tool Mark Examiners' theory at page 86, and the Nichols' article that I just quoted from before at page 589 is defined as follows:

"Agreement is significant when it exceeds the best agreement demonstrated between tool marks known to have been produced by different tools and is consistent with the agreement demonstrated by tool marks known to have been produced by the same tool."

Now, as pointed out at page 560 of *Willock* and acknowledged by the *Montiero* case in the AFTE theory at page 86, this is inherently subjective, and indeed, as we'll see in just a moment, the PCAST report says it's also circular.

When you say that the agreement is significant when it exceeds the best agreement demonstrated between tool marks known to have produced by two different tools, that necessarily encompasses the proposition that different tools can produce marks that are the same. And therefore, in order to take the first step of a sufficient agreement, the tool mark examiner has to have some ability to differentiate the fact that at some

point the number of agreeing marks between the questioned bullet and the source gun are more than you would expect to find from marks that are the same or similar which were produced by firing from different guns. That's the first step, and yet there is no numerical value that allows us to know whether is it one, is it two, is it four, is it 12, is it 17. How do we know that? Definition doesn't answer that.

The same circularity applies to the second half, which is, is it consistent with the agreement demonstrated by tool marks known to have been produced by the same tool. Now, what does that mean? So that means that by definition it can't be avoided that marks, that when you have -- with a tool as a handgun, and the items are being looked at are the cartridges and the bullet fragments, that the very definition of sufficient agreement acknowledges that there will be different marks on the bullets produced by the same tool. And that you expect there to be differences, but when there are similarities, if they are consistent with, whatever that's supposed to mean, the agreement demonstrated by marks known to have been produced by the same tool, then that plus the first component of the definition allows the examiner to express the opinion that there is sufficient agreement and express the opinion, if allowed by the court, that the gun produced was the gun that discharged that ammunition.

This is a subjective test and nothing since 2010 and

nothing before it and nothing after it has done anything to change the subjectivity of this test. As pointed out at page 560 of *Willock*, there is no quantitative standard for how many striations or marks need to match or line up. It's based upon a holistic assessment of what the examiner sees.

And as the *Glynn* case, the *Glynn* court observed, it's inherently vague, and they are significantly subjective. Indeed, Ronald Nichols, who is the -- probably the most well-known ATF agent and AFTE supporter of the substantial agreement and firearms analysis methodology, who is assigned to the ATF bureau in San Francisco, at least was in 2010, as he said in his article previously quoted in the *Willock* article at page 589, there is no universal agreement as to how much correspondence exceeds the best known non-matching situation.

So, how do we try to square the circle? Well, as noted in *Willock*, the -- the AFTE methodology places a great deal of premium on documentation of the reasons concluding that there is a match. And they say that it can include diagrams, photographs or written descriptions, plus what they refer to as peer review, which is not peer review. It's not publication in journals that expose the methodology for others to review and criticize, but rather it just means verification by some other tool mark examiner.

The AFTE standard, and I'm quoting now from the standardization of comparison documentation cited at page 561 of

Willock says: "The case record must contain documentation of the observations that serve as the basis for a reported conclusion."

Now, I'm no longer quoting. It goes on to say there's a lot of latitude, and I now quote, "laboratories are afforded latitude in establishing how this should be accomplished. At a minimum, the documentation must include interpretable depictions or descriptions of the agreement or disagreement of individual and/or class characteristics to the extent that another qualified firearm and tool mark examiner, without the benefit of the evidence itself, can review the case record, understand what was compared, and evaluate why the examiner arrived at the reported conclusion. The case record must clearly describe the label, what items are depicted."

Now, to go back to Rule 702, the -- the essence of trying to apply 702 methodology after the 2000 changes, what guidance does the rule or do the advisory notes give us as to what do you do when the opinion that is to be expressed at trial rests primarily upon the experience and subjective evaluation of the expert.

The Advisory Committee notes say as follows: "If the witness is relying solely or primarily on experience, then the witness must explain how that experience leads to the conclusion reached, why that experience is sufficient basis for the opinion and how that experience is reliably applied to the facts.

The trial court's gatekeeping function requires more than simply taking the expert's word for it and that is critical."

Now, in the *Willock* case, I looked at a number of decisions, all of them by definition before 2010, that have begun to question what was happening with tool mark evidence. It was the Second Circuit's case in 2007, *United States versus Williams*. It noted that *Daubert* did not grandfather or protect from *Daubert's* scrutiny evidence that had previously been admitted.

And there was the *Green* case that acknowledged that courts were obliged to critically evaluate tool mark and ballistic evidence even though it had been accepted for years, 405 F.Supp. 2d at 104. And I acknowledge that beginning in 2005, in the National Research Council's Forensic Science Report and thereafter with the National Research, the NRC Forensic Science Report that the scientific community as charged by the National Academies of Science had begun to look at forensic analysis and tried to determine whether or not it met reliability standards after *Daubert* in 702.

And in the National Research Council's Forensic Science Report in 2005, as quoted at page 565 of *Willock*, I noted that the NRC Forensic Science Report concluded, "because not enough is known about the variabilities among individual tools and guns, we are not able to specify how many points of

similarity are necessary for a given level of confidence in the result. Sufficient studies have not been done to understand the reliability and repeatability of these methods.

The committee agrees that class characteristics are helpful in narrowing the pool of tools that may have left a distinctive mark. Individual patterns from manufacture or from where might in some cases be distinctive enough to suggest one particular source, but additional studies should be performed to make the process of individualization more precise and repeatable."

Now, what the NRC was saying is that the validity of the tool mark analysis methodology depends upon establishing two things. The uniqueness of individual marks and the reproducibility of those unique marks on ammunition when it is fired from the firearm. And unless those two things can be established, then you have not met the requirement of showing that the methods and principles used are reliable.

The point that was also made in the *Willock* case which is important was that the NRC ballistic imaging report cited a article by a Biasotti, B-I-A-S-O-T-T-I, published in 1959 which said, and I quote now, this is at page 72 of the *Willock* decision, "The average percent match for bullets from the same gun is low, and the percent match for bullets from different guns is high."

So what that means is, is that when you look at all

the marks on bullets known to have been fired from the same gun, you can focus in on matches that, on marks that match, but there's a large number of marks that don't match. And when you focus on cartridges fired from different guns, there's a large number of similar marks as compared to distinct marks.

So the question is, if you are applying an inherently suggestive approach of sufficient agreement and you know that marks, there are going to be a relatively large number of marks on cartridges fired by different firearms, and when you compare all of the marks on cartridges fired by the same firearm, there will be a lot that don't match. How do you determine sufficient agreement? And there is nothing, at least at that time in the methodology of tool mark examination that gave you the answer to that question, other than just simply, well, this is the experience of the examiner.

Since 2010, the debate on forensic evidence has not abated. It's increased. In September of 2016, the report to the president from the forensic science on forensic science in criminal courts from the President's Advisory Council on Science and Technology came out. This is referred to as the PCAST.

The report indicated that it was provided after the analysis by the National Research Council in 2009 and was designed to bring forward analysis of forensic comparison method analysis. So feature comparison methods included things such as single sourced DNA analysis, complex mixture DNA analysis, bite

mark analysis, latent fingerprint analysis, firearms analysis, footwear analysis, hair analysis, and this report was -- made recommendations to the National Institute for Standards and Technology, NIST, the Department of Commerce, recommendation for the FBI laboratory, recommendations to the attorney general, recommendations to the judiciary.

And I want to refer now to the portions of the PCAST report that focused on firearms analysis which advances five years to the most immediate past. And before doing that, I want to make one further point on *Daubert* that I didn't make earlier. The result of *Daubert* analysis in theory, anyway, is not to analyze the correctness of the decision reached, but rather to analyze the reliability and methodology and sufficiency of the factual data.

So, what do we know from the 2016 PCAST report? And I might add also that the PCAST folks that were -- the co-chairs were the assistance of the president of the United States from science and technology, Professor Eric Lander, president of the broad institute of Harvard and MIT. The vice chairs included professors from computer science and integrated biology from the University of Texas. Members included the president of large corporations, professors from astrophysical science and international affairs from Princeton, physics professors from the University of Maryland, professors from the school of Natural Resources and Environment from the University of

Michigan. Natural economics, University of Maryland, the Kaiser Permanente School of Medicine, the planning dean and professors from various fields of science and technology across the group. The PCAST working group, the president was Professor Lander, Professor Gates from the Center for String and Particle Theory, Professor of Physics, University of Maryland, Professor Chin, distinguished professor emeritus and electrical engineering and computer science from Berkeley.

The Professor Press from the Computer Science and Integrated Biology Department at the University of Texas, Austin. Professor Schrag, professor of environmental science and director of Harvard University Center for the Environment. The staff included Diana Pankedich P-A-N-K-E-D-I-C-H, from the American Association for the Advancement of Science and Technology, and Tianna Simoncelli, S-I-M-O-N-C-E-C-L-L-I, senior advisor to the director of the Broad Institute at Harvard, and Kristin Sorelli, advisor to public policy and special projects at Harvard and MIT. And they had a number of judges and academics who also were on the panel.

Beginning at page 104 of the PCAST report it began its analysis of the methodology for firearms analysis. And it started with the notion that, "Firearms analysis," quote, "is based on the idea that tool marks produced by different firearms very substantially enough owing to variations in manufacture and use to allow components of fired cartridges to be identified

with particular firearms.

"For example, examiners may compare questioned cartridge cases from a gun recovered from a crime scene to test fires from the suspected gun." Talks about class characteristics and individual characteristics, and it made the following comment at page 104.

"PCAST expressed concerns about certain foundational documents underlying the scientific discipline of firearm and tool mark examination. In particular, we observed AFTE's theory of identification as it relates to tool marks which defines criteria for making an identification is circular. The theory states that an examiner may conclude that two items have a common origin if their marks are in sufficient agreement, where sufficient agreement is defined as the examiner being convinced that the items are extremely unlikely to have a different origin. In addition, the theory explicitly states the conclusions are subjective."

The report continues at page 105 and noted that there had been a lot of testimony -- a lot of attention in the recent past to try and to prove the theory that every gun produces unique tool marks. It made reference to the 2008 NRC report, and that report found, and I quote from page 105 of the PCAST report, "The validity of the fundamental assumptions of uniqueness and reproducibility of firearms related tool marks had not yet been demonstrated, and given current comparison

methods, a database search would likely return too large a subset of candidate matches to be practically useful for investigative purposes."

The report went on to say that, of course, it's not necessary that tool marks be unique for them to provide useful information about whether a bullet may have been fired from a particular gun, but, quote, "However, it is essential that the accuracy of the method for comparing them to be known to be based upon empirical studies."

They then cited from the 2009 NRC report, they concluded that there had not been sufficient studies done yet, and the footnotes to this analysis make reference to many of the studies that they looked at.

We heard a lot during the hearing yesterday and today about various studies that were attached to the exhibits that had been submitted to me and that I have reviewed. These various studies are discussed at the PCAST report beginning at page 106, and they differentiated between non-black box studies, which are referred to as set based analyses. And set based analyses essentially means that the examiners are aware that they have within the sample bullets, portions known to have been fired by the gun in question when in real life situations, crime scene analysts don't know that. They're seeking to determine whether that's the case, at least oftentimes.

It went on to say that, "Because firearms analysis is

at present a subjective feature comparison method, its foundational validity can only be established through multiple independent black box studies."

They concluded that, "Although firearms analysis has been used for many decades, only relatively recently has its validity been subject to meaningful empirical testing. Over the last 15 years the field has undertaken a number of studies to establish the accuracy."

It goes on to say that, "While the results demonstrated that the examiners can under some circumstances identify the source of fired ammunition, many of the studies were not appropriate for assessing scientific validity and estimating the reliability because they employed artificial designs that differ in important ways from problems faced in casework."

The analysis continued: "Specifically, many of the studies employed," quote, "set based analyses in which the examiners are asked to perform all peer-wise comparison within or between small sample sets."

They go on to say that, "The study design in some of these tests had serious flaws because the comparisons are not independent of one another, but rather involve internal dependencies that inform the examiners' answers and sometimes can allow the examiners to make inferences about the study design."

They conclude that, "Set based studies are not appropriately designed. Black box studies from which you may determine empirical studies of validity."

They went on to talk about certain types of set based comparisons that they referred to as within set comparisons and set-to-set comparisons and concluded that when they studied four specific such cases to include the Fadul, F-A-D-U-L, study, which is referred to here, that dealt with ten consecutively manufactured firearm slides, "examiners were given a collection of questioned bullets and/or cartridge cases fired from a small number of consecutively manufactured firearms of the same make. A collection of bullets or casings known to have been fired from these same guns. They were asked to perform matching exercises."

It says, "This closed set design is simpler than the problem encountered in casework because the correct answer is always present in the collection. In such studies examiners can perform perfectly if they simply match each bullet to the standard that's the closest. By contrast in an open set study, as in casework, there's no guarantee that the correct source is present. Closed set comparisons would thus be expected to underestimate the false positive rate."

They concluded on page 109, "In short, the closed set design is problematic in principal and appears to underestimate the false positive rate in practice. The design is not

appropriate for assessing scientific validity and measuring reliability."

They then at page 109 talked about the set-to-set comparison in a partially open set, the Miami/Dade study and then talked about recent black box study analysis that had taken place, referring to the Ames Laboratory study, which was introduced in this hearing.

And that actually did employ a design that had black box configuration, and in that particular study they found that when 218 examiners did their examination, one of whom was Mr. McVeigh, that they -- that there were 2,178 different source comparisons. There were 1,421 eliminations, 735 inconclusives and 22 false positives. The conclusive rate was 33.7 percent, and the false positive among conclusive examinations was 1.5 percent with a confidence interval of 2.2 percent.

The false positive rate corresponds to an estimated error rate of one in 66, so the false positive corresponds to an error rate of one in 66. This is the best test done so far. Every 66 cases, there will be one error with an upper bound being one in 46; that would be the highest error rate, so that's the range.

The PCAST report concluded, "The results for the various studies are shown in table two. The table show a striking difference between the closed set studies where a matching standard is always present by design, and the

non-closed studies where there is no guarantee that any of the known standards match.

Specifically the closed set studies show a dramatically lower rate of inclusive examination and false positives. With this unusual design, the closed set studies that have been criticized, in other words, examiners succeeded in answering all the questions and achieved essentially perfect scores. In more realistic open designs, these rates of false positives are much higher."

The conclusions they reached, "The early studies indicate that examiners can under some circumstances associate ammunition with the gun from which it was fired. However, as described above, most of these studies involve designs that are not appropriate for assessing scientific validity or estimating the reliability of a method as practiced. Indeed, comparison of the study suggests that because of their design, many frequently cited studies seriously underestimate the false positive rate."

The conclusion continues at page 111. "At present there is only a single study that was appropriately designed to test foundational validity and estimate reliability, the Ames study. Importantly, the study was conducted by an independent group unaffiliated with a crime laboratory."

It goes to say that, "The scientific criteria for foundation of validity require appropriate design studies by more than one group to ensure reproducibility. Because there

has only been a single appropriate design study, the current evidence falls short of the scientific criteria for foundation of validity," and concluded that there was a need for additional analysis."

So, they summed it up in a chart on page 112. Foundational validity, quote, "PCAST finds that firearms analysis currently falls short of the criteria for foundational validity because there's only a single appropriate design study to measure validity and estimate reliability. Scientific criteria for a foundation of validity require more than one such study to demonstrate reproducibility.

Whether firearms analysis should be deemed admissible based on current evidence is a decision that belongs to the courts. If firearms analysis is allowed in court, the scientific criteria for validity as applied should be understood to require clearly reporting the error rates as seen in appropriately designed black box studies estimated to be at one in 66 with a 95 percent confidence limit of one in 46 in the one such study to date."

Now, as the government has properly pointed out, not everybody accepts the PCAST report. Indeed, there was a rather significant unfavorable response by a number of organizations to include the Department of Justice.

I'm going to turn to some of those in a moment, but first I want to cite from Defense Exhibit 1, the Ames study.

And the Ames study, again, the essence of that is found on page 3 of the study, and it was summarized as follows.

First of all, it's done by Ames Laboratory associated with Iowa State University, so it was a -- an independent laboratory. It was not associated with the Association of Firearm and Tool Marks Examiners, which is the association of these folks who come into court and testify about firearms.

Here's the summary. "Responses were received from 218 participating examiners. The rate of false negatives was quite low with the error distributed across examiners with various backgrounds, state, federal, local, private. The overall rate of false positives estimated as 1.01 percent from comparisons known to be from different firearms, but reported as identifications was significantly higher.

"However, most of the error rates were reported by a small number of examiners, that is, individual examiners have varying error rates. For most examiners this is quite low while for some it is relatively high. Hence, the overall rate is best interpreted as an average of widely varying individual rates. Inconclusive results were not recorded as errors. Rates of poor quality marked production for these handguns varied. False positive and false negative error rates for individual examiner performance were measured. The rates were not uniform across the sample population with a few examiners providing most of the false positives."

Now, the significance of that report is that with 218 examiners, the actual false positive rates found were not uniformly disbursed among all those examiners. Some were very good and some were not. So how do we decide those that are good and those that are not?

Well, one way requires strict adherence to the methodology and documentation of how the methodology was applied in that particular case so that looking at just the records and the documentation without looking at the evidence itself, you can determine whether or not the conclusion reached was a proper conclusion.

Now, there were a number of other organizations including OSAC, the Department of Justice, individual tool mark practitioners, that went back to PCAST and said, you got it wrong. There are studies you didn't look at. You blew off these other studies that meet the criteria that you say we should have done, and you've, you know, cast aspersions on us that were not justified.

So PCAST considered those and came out with an addendum that they authorized just a little more than a year ago, January 6th, 2017, found at Exhibit C to the defense motion. It sort of summarized where PCAST left off to explain what it did in response to the criticism from many sources as to what it had done.

It said at page 1, "In this report, PCAST noted that

the only way to establish scientific validity and degree of reliability of a subjective forensic feature comparison method, that is, one involving significant human judgment, is to test it empirically by seeing how often examiners actually get the right answer. Such an empirical test of a subjective forensic feature comparison method is referred to as a black box test. The point reflects the central tenant underlying all science. An empirical claim cannot be considered scientifically valid until it has been empirically tested."

It goes on to say that, "Practitioners of a subjective forensic feature comparison method claim that through a procedure involving substantial human judgment they can determine with reasonable accuracy whether a particular type of evidence came from a particular source, the claim cannot be considered scientifically valid and reliable until one has tested it by, one, providing adequate number of examiners with an adequate number of test problems that resemble those found in forensic practice; and two, determined whether they get the right answer with acceptable frequency for the intended applications.

"While scientists may debate the precise design of the study, there is no room for debate, but the absolute requirement for empirical testing."

Now, having said that, we stated what their ground was from before, they summarized the responses that they had gotten.

Page 2 of the addendum.

The following reports released, "PCAST received input from stakeholders expressing a wide range of opinions. Some of the commentators raised the question as to whether empirical evidence is truly needed. The Federal Bureau of Investigation, which clearly recognizes the need for empirical evidence and has been a leader in performing empirical studies in latent print examination, raised a different issue. Specifically, although PCAST has received detailed input on forensic methods from forensic scientists at the FBI laboratory, the agency suggested that PCAST may have failed to take account of some relevant empirical studies.

A statement issued by the Department of Justice on September 20th, 2016, the same day the reports released opined that, quote, "The report does not mention numerous published research studies which seem to meet PCAST criteria for appropriately designed studies providing support for a foundation of validity. That omission discredits the PCAST report as a thorough evaluation of scientific validity."

So, what was the response of PCAST? They say, quote, "Given its respect for the FBI, PCAST undertook a further review of the scientific literature and invited a variety of stakeholders, including the DOJ, to identify any published appropriately designed studies that had not been considered by PCAST that established the validity and reliability of any of

the forensic featured comparison methods that PCAST report found to lack such support. As noted below, DOJ ultimately concluded that it had no additional studies for PCAST to consider.

"However, PCAST did receive written responses from 26 parties, including federal agencies, forensic science and law enforcement organizations, individual forensic science practitioners, a testing service provider, and others in the United States and abroad. It noted that many of the responses were extensive, detailed, thoughtful, and covered a wide range of topics."

It acknowledged its gratefulness for the time that those folks took who opined on this important topic.

So based upon what they had received, they provided the following analysis: "While forensic science organizations agreed with the value of empirical tests of suggested forensic feature comparison methods, that is, black box test, many suggested the validity and reliability of such a method could be established without actually empirically testing the method in an appropriate setting. However, PCAST noted notably, however, none of these respondents identified any alternative approach that could establish the validity and reliability of a subjective forensic feature comparison method."

It went on to say that after reviewing what they had received, quote, "There remains confusion as to whether these elements can suffice to establish validity and degree of

reliability of a particular method."

Subparagraph 1: "Forensic science literature contains many papers describing variation among features. In some cases the papers argue patterns are unique." It goes on, "Such studies can provide a valuable starting point for a discipline because they suggest that it may be worthwhile to attempt to develop reliable methods to identify the source of the sample based on feature comparison. However, such studies, no matter how extensive, can never establish the validity or degree of reliability of any particular method. Only empirical testing can do so.

Conclusion two: "Forensic scientists rightly cite examiners' experience and judgment as important elements in their disciplines. PCAST has great respect for the value of examiners' experience and judgment. They are critical factors in ensuring that a scientific eval and a reliable method is practiced correctly. However, experience and judgment alone, no matter how great, can never establish validity or degree of reliability of any particular method, only empirical testing can do so."

It then responded to input it received from the Organization of Scientific Area Committee's Friction Subcommittee, that's the OSAC, we heard testimony about that yesterday. The report at page 4 says: "In its response to PCAST's call for further input, the Organization of Scientific

Area Committee's Friction Ridge Subcommittee, whose purview includes latent fingerprints, raised some very important issues."

They said, "While the OSAC FRS agrees with the need for black box studies to evaluate overall validity, they expressed concern that that view could unintentionally stifle future research aimed at dissecting components of the black box to transition it from subjective to objective method. As for the friction ridge discipline, which is different from the tool marks discipline, PCAST concluded that it applauds the works of friction ridge discipline which has set an excellent example by undertaking both, one, path breaking black box studies to establish validity and equally of reliability of latent fingerprint analysis; and two, insightful like box studies that shed light on how latent print analysts carry out their examination, including forthrightly identifying problems and needs for improvements."

It goes on to say, however, the situation is different for subjective methods whose validity and degree of reliability has not been established by appropriate empirical studies. If a discipline wishes to offer testimony based on the subjective method, it must first establish the method's validity and degree of reliability which can only be done through empirical studies."

They then analyzed the challenges to the completeness

of PCAST evaluation, page 5, and they said the following: "Finally, we considered the important question raised by the DOJ in September of whether a PCAST failed to consider numerous published research studies which seemed to meet PCAST criteria."

The analysis goes on: "PCAST reexamined the five methods evaluated in its report for which the validity and degree of reliability had not been fully established. We considered the more than 400 papers cited by the 26 respondents, the vast majority had already been reviewed by PCAST in the course of the previous study at the suggestion of John Butler at the National Institute of Standards and Technology, NIST. We also consulted Interpol's extensive summary of the forensic literature to identify additional potentially relevant papers. Although our inquiry was undertaken in response to the DOJ's concern, DOJ informed PCAST in late December that it had no additional studies for PCAST to consider. It then applied its ultimate amended conclusions to various disciplines."

Page 6, it talks about firearms analysis, and it concluded in its amended report -- addendum to its report. "In its report PCAST reviewed a substantial set of empirical studies that had been published over the past 15 years and discussed a representative subset in detail. We focus on the ability to associate ammunition not with a class of guns, but with the specific gun within the class. Firearms discipline clearly recognizes the importance of empirical studies. However, most

studies use flawed designs.

"As described in the PCAST report, set based approaches can inflate examiners' performance by allowing them to take advantage of internal dependencies in the data. The most extreme example is the closed set design in which the correct source of each question sample is always present. Studies using the closed set design have underestimated the false positive and inclusive rates by more than one hundredfold. This striking discrepancy seriously undermines the validity of the results and underscores the need to test methods under appropriate conditions.

"Other set based designs also involve internal dependencies that provide hints to examiners, although not the same extent as closed set designs. To date, there has been only one appropriate designed black box study, a 2014 study commissioned by the Defense Forensic Science Center and conducted by the Ames Laboratory, which reported upper 95 percent confidence bound on the false positive rate of 2.2 percent.

"Several respondents wrote to PCAST concerning firearms analysis. None cited additional appropriate designed black box studies similar to the Ames Laboratory study."

Went on to conclude, "The Organization of Scientific Area Committee's Firearms and Tool Mark Subcommittee took the more extreme position that all set based designs are appropriate

and that they reflected actual casework because examiners often start their examinations by sorting sets of ammunition from a crime scene.

"OSAC's FTS argument is unconvincing because, one, it fails to recognize that the results from certain set based designs are wildly inconsistent with those from appropriately designed black box studies; and two, the key conclusions presented in court do not concern the ability to sort collections of ammunition as tested by a set based designs, but rather the ability to accurately associate ammunition of a specific gun as tested by appropriately designed black box studies."

It went on to say as follows: "Courts deciding on the admissibility of firearms analysis should consider the following scientific issues: One, there's only a single appropriate black box study employing a design that cannot provide hints to examiners. The upper confidence bound on the false positive rate is equivalent to an error rate of one in 46.

"Two: A number of older studies involved the seriously flawed closed set design, which has dramatically underestimated the error rates. These studies did not provide useful information about the actual reliability of firearms analysis.

"Three: There are several studies involving other kinds of set based designs. These designs also include," I'm

sorry. "These designs also involve internal dependencies that can provide hints to examiners, although not to the same extent the closed set designs do. The large Miami/Dade study cited in the PCAST report and the small study cited by Bunch fall into those categories.

These two studies have upper confidence bounds corresponding to error rates in the range of one in 20 from a scientific standpoint, scientific validity should require at least two proper design studies to ensure reproducibility. The issue for judges is whether one poorly designed study together with ancillary evidence from imperfect studies adequately satisfies the legal criteria for scientific validity. Whatever the courts decide, it's essential the information about error rates is properly reported."

So that's the -- that's the PCAST in its wake.

Now let's turn to this particular case because that general background is important in terms of my cast under Rule 104(a) when deciding whether or not Mr. McVeigh can testify, and if so, as to what.

First of all, as to his qualifications. There has really not been much challenge to his qualifications. I will say again on this record that I've been a judge for 21 years, and I've been a trial lawyer for almost 50. And I have seldom seen an expert as candid, forthright, acting without a shred of being defensive, trying his hardest to understand questions,

respond without being argumentative.

It appears as though Prince George's County and the people of Maryland are lucky to have a person of his integrity and dedication working for the crime lab in Prince George's County. There is no challenge to his qualifications. Although he's not a member of the AFTE, he took their instruction, he's been qualified by his lab, and he has done the examination in this case based upon what he has identified in terms of his training and experience. So I find easily that he has the knowledge, training, experience, background, education and skill to testify in the form of an opinion or otherwise.

The next thing I have to talk about is whether he has sufficient data to base his test. The challenge really hasn't been to the sufficiency of the data, but rather to the reliability or the methodology and whether it was reliably applied.

In doing that, I want to start with his examination report. This is Exhibit 5 to ECF number 78. It's a series of documents, the report itself, that begin at Bates number JLM OO773 through really 782, are the guts of it, and I'm going to go through those briefly now.

Page 773 reports the result of examination. He says there are five fired cartridges. I'm leaving portions of it out. They were microscopically intercompared to test fire exemplars from the Rock Island Armory pistol recovered from the

incident.

Quote: "The referenced items from incident A were identified as having been fired in from the referenced pistol from incident B." And there are no further grafts or anything else, that's what I would call face sheet of the conclusion sheet.

Page 774 is a result of the examination. He says there are five fired cartridge cases, PW 1 through PW 5, and two bullet items, PW 6 and PW 7, recovered in incident A that were microscopically intercompared to test fire exemplars from the Rock Island Armory brand pistol. And there's some more non-opinion information, and it then concludes the referenced items from incident A were identified as having been fired in from reference pistol from B.

And below it there's a series of initials. It says, FCC's, which he testified were fired cartridge cases. BFI, breach face impression, which he testified were marks on the rim, and marks at the rim. And he just identifies those. So presumably you can infer from that that that's what he looked at when he was reaching his conclusion. Bullets, and then it says, L-I-M-P-S, which stands for land impressions.

Now, significantly, although this does explain what his opinion is, gun -- incident B gun produced was the gun from which incident A bullet portions were fired from. What he doesn't do is provide the analysis of how he reasoned from the

look at the bullet fragments to these conclusions. We know that he looked at fired cartridge cases. But that doesn't tell us a lot. He looked at the breach face impressions, that narrows it down, and marks on the rim, which further narrows it down.

But there are marks not only on the rim, but on the cylinder part of the cartridge, and we know, as I quoted from *Willock*, that from publications going back to the 1950's, there is a large number of similar marks on items known to have been produced by different firearms. And a equally large number of marks on bullet components fired from the same firearm that are not similar.

And what we don't know is from this face sheet or from the photographs is when you line up a portion of the firearm component looked at, and you look to see lines that seem to come together and lines that don't, which ones were the ones that he relied upon in reaching his conclusion that there was sufficient agreement. Take a look at 775.

First of all, the one on the left is at 36 magnification. It's at the rear of the cartridge, and it is a portion that looks like about 25 percent of what the actual cartridge, a back look like. If you take a look at it, there are lines that sort of seem to go from -- and there's about a 40 percent portion of the left-hand side of the picture that's from the one sample, and on the other side about 60 percent from the other. And collectively put together they look to be about

roughly 25 or so percent of the back of a bullet that you would look at if you were holding it up and looking at where the firing pin connects the rim.

Some lines start from the one sample and go all the way through to the other, but others start on one and don't go all the way through so it's possible to look at this and say, well, those are ones that don't match up, but these are ones that do match up, and some of the ones that match up are at the top and at the bottom and sort of in the middle part a little bit here and there, but others that do not.

Now, how do I know that the numbers that match are consistent with the numbers that match on cartridges known to be fired from the same gun, and that the numbers that don't match are less than the numbers that are known not to match from those fired by different guns.

If you look to the picture on the right, it's even a smaller portion of the cartridge. It's a very small part of the circular rim. In fact, it doesn't even look circular. It looks sort of like an almond shaped component that is obviously just a very small portion. It's also at a much higher -- it's at 56 level of magnification. Again, if you look at the part on the left, there's some parts that seem to go clear through to the one part on the right, but there's some other parts that don't seem to do that.

And then there's some parts below that are dark and

you can't really tell what they show. Again, how do I know how you find that there were enough similarities to be consistent with similarities from bullets fired from the same gun and greater than the number of similarities known to exist in bullets fired from different guns? There's no way I know that, nor does the report explain that to me.

The same can be said at page 776. At 777, there are conclusions. The same one fired from the same unknown firearm. There's no photographs here. It does talk about what items were looked at, but again, it doesn't provide any explanation.

You have on page 778 the fired cartridge case worksheet. That, of course, provides data regarding the kind of ammunition. And it says notes, the micro shows it was identified as having been fired by the same unknown firearm, but it doesn't explain how this data supports that conclusion. Same with 779, same with 780, same with 781.

On 782, we had certain measurements that were taken that were fed into the NIBIN system, and it resulted in a reporting of the other types of firearms that could produce those same characteristics. And there were a total of 33 manufacturers listed that did not include the one that we know manufactured the firearm.

Now, I'm not -- it appears as though what Mr. McVeigh did is consistent with the internal procedures from his lab, and indeed, when his lab has been audited by others who do this kind

of lab work, they've said that your paperwork is sufficient.

It was interesting that he testified that when he was reviewed, what the reviewer did was he would frequently just have the microscope up and say, hey, Bill or Mary or whatever, come on over and take a look at this. They would look at the same microscope at the actual evidence, make their own determination and say thumbs up or thumbs down.

But that's not what the standard from the AFTE says, and that's not what the advisory note to Rule 702 says. It says that, "When you're basing it upon a subjective analysis and experience, the actual notes have to tell you how you applied the methodology to the facts and made that result," and simply saying, well, I don't tell out the one, the number of marks on there because there were so many, I just line them up and I can look at them, in my mind that does not meet what the Advisory Committee required, it does not provide enough information to show how the methodology was applied in this particular case.

Similarly, the peer review, so to speak, is reviewed by another examiner from the same office. There was also one from, I think the District of Columbia, whose initials were put on there. We don't have any separate reports from them, and what it appears as though is that they didn't look at just the report itself and go back to see whether they agreed with the results, they rather looked at the ammunition. And that, that is in my mind the concern that I have about the methodology and

the application of the methodology to the facts.

So where does this leave me? The one thing that I said in the *Willock* decision, my recommendation to Judge Quarles and what I repeat now is that as PCAST sort of recognizes, they're looking at this area through the lens of science, and it's pretty clear that they say that for this kind of subjective analysis, you must have empirical research, it's got to be black box, you got to have the right error rates, the design has got to be such that it mirrors real life. That makes all the sense in the world to me for science, but they drew the line by not telling judges, this is what you should do, and they made that very clear throughout.

So what I'm left with is, with what I have from an admittedly qualified and dedicated examiner in a case, is there anything about this evidence that will pass muster from a point of view of 702. And here I come back to one of the questions I asked Mr. Oppenheimer because if we didn't have any expert at all, but we knew that there was a .45 that was the caliber of the handgun, and a .45 is found with the defendant, and if we knew that the cartridges found at the crime scene were Winchester and that the cartridges found in the handgun were Winchester, and if we knew that there were other similarities, juries have throughout the history of this country been able to independently determine the authenticity under Rule 901(b)(3) by looking at known samples and unknown samples and deciding for

themselves whether or not they were from the same source. That's an accepted way of authentication.

And certainly would it be helpful for the jury to know that I've already ruled that the .45 comes in as being something taken from the possession of the defendant, he pleaded guilty to having it in D.C. Superior Court, and the search was a valid search, I found. That was objected to, but I overruled that objection.

So we know that the gun was found with the defendant, and we know that we have bullets that were found from the crime scene, we know we have bullets that were found from the gun. And I don't see any reason why it would not be helpful to the jury for Mr. McVeigh to testify with his photographs and matching up the marks that he saw that were similar and pointing out the characteristics that were similar between the firearm cartridges fired at the scene and what was test fired, subject to cross-examination, subject to Evidence Rule 803(18), the learned treatise hearsay exception, where certainly a defense can read into the record, and 803(18) says, "To the extent relied upon by an expert during direct examination or called to the attention on cross-examination, learned treatises," which I find the PCAST reports would be, "can be read into the record." And whether or not Mr. McVeigh acknowledged them or not, that's evidence that comes in and can be read to the jury so they can understand.

In addition, certainly Mr. Nixon can testify at trial, if that's what the defense wants him to do.

In addition, I believe that it would be appropriate for Mr. McVeigh to be able to express an opinion that the marks that were produced by the -- that were found on the crime scene cartridges are consistent with the marks that were found on the test fire from the .45 known to be associated with the defendant, but I will not permit him to express the opinion that they were fired by the same gun, and I will not permit him to express any confidence level as to it.

So -- so we're here clear on what I'm ruling. Mr. McVeigh can testify. Now, his qualifications are subject to being brought out before the jury by the government and challenged or cross-examined by the defense. He can talk about what he did and what he tested and what he looked at. He can put up his pictures. He can show the similarities between one and the other, and he can even express the opinion that the marks from the .45 that he test fired, the marks on the cartridges are consistent with the marks on the other one that were found at the crime scene, but I won't allow him to express the opinion that they were produced by the same gun, and I won't allow him to express a confidence level as to his opinion. That's as far as his opinion can go.

The defense will be able to cross-examine, offer under 803(18) information as learned treatises. On rebuttal, the

government may offer any of the studies that they thought were helpful under 803(18) if they want to do that. Be aware, counsel, that when you're doing that, the jury doesn't get to see them.

Probably you've been -- you either have appeared to stay awake or have resisted the sopopheric [ph] soporific, excuse me, effect of my language, reading to you from all these different studies, but be aware that the jury may not have the same intestinal fortitude, so I would keep my readings focused, but that is what I believe is appropriate. And why do I say that?

While I am informed by the PCAST analysis, and I think that from the perspective of science they raise legitimate issues, I'm not prepared to say that you got to throw the baby out with the bath water, particularly when you have a dedicated and honest broker like Mr. McVeigh, who is straightforward in what he did and can point out why he drew the observations that he did from the test samples and from the firearm exemplars and properly restrict it in terms of the scope of what he can tell the jury, that that would be helpful to the jury because the jury itself can look at known and unknown samples and decide whether it came from the same gun under Rule 901(b)(3). And so that's something that's been allowed by the Rules of Evidence independently of Rule 702.

Secondly, I have to note that not every court in the

United States has followed what PCAST said. Government cited a case from the Northern District of California that said, well, we don't care what PCAST said, we're going to let it in because it's always been let in. Respectfully, I disagree. The judge's job would be easy if all we could do is say courts have always done, I'm going to do it again. That's where I was in 2010, I'm not there anymore.

2010, I observed that despite growing concern by the courts, no American court that I was aware of had totally excluded the evidence, at most they had restricted what the expert can testify, which is consistent with my ruling today, but you can't ignore what the issues were that were raised by PCAST, and I can say that it takes some time sometimes for the courts to come to bear and to take knowledge of what's going on, and the mere fact that twice out of 80 times Mr. McVeigh has had to go through this exercise is evidence of the fact that most courts do not do it. They don't have the circumstance where they have defense counsel that have the resources and the dedication to bring the issues, and they can't find an expert to testify and provide them with help, and so the evidence rolls in.

The PCAST report is not the end of the story, but it does show that there are issues associated with the validity of the underlying subjective assumption that this science rests on. I'm convinced that the similarities of the characteristics and

the scope of what I've allowed Mr. McVeigh to testify to is sufficiently helpful, does pass muster under, not only 702, but other rules as well, and that with experienced and focused cross-examination from lawyers who have clearly prepared themselves for the task, that the weight that can be given to the evidence I permit will be a point for argument with the jury.

Now, Mr. Oppenheimer, if you want to propose a jury instruction, I'll take a look at it. Whether this is appropriately handled by a jury instruction or by a limiting instruction when the evidence comes in, we can deal with later, but for right now I understand that you've suggested that. I haven't seen it, and we'll have plenty of time during the trial to deal with post-jury instructions. So that's by ruling on this motion.

I had hoped to be able to write an opinion, but with the other work that I have and the other work that we have in this case, I'm likely not to get to it so this transcript will be my opinion. I may come up with a very brief one pager that says for the reasons I stated on the record which are adopted and incorporated by reference, this is my ruling, but my ruling should be fairly clear, I hope, to both the United States and Mr. Oppenheimer. Are there any questions about what my ruling is?

(Pause.)

MS. KAPLAN: No, Your Honor, we're just taking a moment to make sure we --

THE COURT: No, that's all right, take whatever time you want. You got to live by it, you might as well make sure you understand what it is.

(Pause.)

THE COURT: While the government is looking at its notes, are there any questions from defense in terms of my ruling?

MR. OPPENHEIMER: If I could have one more moment here, Your Honor. I just want to make sure I --

THE COURT: Yes.

(Pause.)

(Brief recess.)

THE COURT: All right. Any questions from the United States regarding my opinion?

MR. NAUVEL: Yes, Your Honor. A couple brief questions. I think you've gotten to know just a little bit, Mr. McVeigh, over the past couple of days, and he's very concerned with respecting your order and following it to the letter. But he's also raised concerns also which I'll relay to you, which is that he is concerned with his ethical obligations outside of the testifying piece, outside of the courtroom, and more specifically, AFTE provides a range of conditions, a range of conclusions, and three conclusions I think you heard him testify

to were identified, eliminated, inclusive. And he's also concerned about his protocols, which make similar type of points, so all this to say, the question is, is it consistent with your ruling to say --

THE COURT: Well, let me see if I can -- he can certainly say he did not eliminate this gun. He can certainly say that his -- his views are not -- well, I would say this: He is -- he has not eliminated his gun as having made it. I think if he says that he did reach a position that it was not inconclusive, then I don't want to get close to what I said he can't go beyond. And he can certainly say that he did not eliminate this gun as being a possible source to it, but that the best he can say is that, you know, and if you want him to say that I, you know, something to the effect of that the Court has instructed him to, you know, that he can say that it's consistent with or you can frame the question so that he is responding directly to your question.

I don't want to -- I certainly understand that he made different opinions I'm not letting him put in, and if he is faithful to my ruling, it's not because he's done anything wrong, it's because I've ordered it, and he's complying with the Court Order. I think the way it can be handled is, is that, you know, it can be done in a question, do you have an opinion that you can express to the jury as to whether the marks on the bullets fired by the gun are consistent with the marks found on

the other one. And then he can say obviously he thinks it's consistent, that's part of what his opinion was, it's just not expressed in the terminology that he's accustomed to using.

So he's nodding his head, I think he understands that. I think the best way to handle it is by a question posed by the government to allow him to make that specific response.

MR. NAUVEL: Could he say that they are -- the -- that the marks have been identified, have been identified as being consistent with the marks on the test fire?

THE COURT: In his opinion, the marks on the test fire are consistent with the marks he found -- with these marks that he found on the exemplars.

Now, the defense can say, well, it's not consistent with this one, that one and the other one, but you can say it that way, yes. He just can't say it's from the same gun, and he can't say his confidence level.

MS. KAPLAN: So, to the extent the word "identify" has particular meaning within his field, and, Mr. McVeigh, if you would come right up here in case I lose it, you can whack me on the shoulder.

So I think what he asked us is can he say that in his opinion the marks on the bullets fired from the gun have been identified as being consistent with the marks on the bullets recovered from the crime scene. I don't know that that is meaningfully different from what Your Honor has suggested, it

just includes that would identify, which I think is important for your working with.

THE COURT: Yeah, explain what your concern is, sir.

MR. McVEIGH: Oh, just the points between the two points, what counsel just said regarding identified as having consistent with or is sufficient to replace the word "identified" with the phrase "consistent with" to honor your ruling.

THE COURT: Yeah, for my ruling I'm substituting the word "identify," which is a word of art that you -- when you say identify, it's the same thing as saying in my opinion it's the same gun, right? That's what it means. I'm not going to allow you to say that, but you can say that -- and you can point to the marks and say these marks here, here, here, here, here, are consistent with marks made from the gun.

MR. McVEIGH: That's the clarification.

THE COURT: And that's what I'll allow you to do, and you can point out however many of them. You can say there's so many of them, there's just this, this and this, but he gets to cross-examine on the ones that don't match, and he can cross-examine and read into the record the, you know, all the studies and everything else that he wants, and he can call his own expert to say no, it's not. So that's what I'm going to allow to have happen.

MR. McVEIGH: Thank you, Your Honor.

THE COURT: Anything further from the government?

MS. KAPLAN: And so for him to say the marks on the bullets recovered from the crime scene are consistent with the marks made from the gun would be an appropriate way for him to --

THE COURT: Are consistent with these marks, particular marks. I'm going to allow him to identify which marks he says are consistent, not just a general thing that are consistent with the same one. Because that's seems to be the ultimate conclusion, I'm not going to allow him to say.

MS. KAPLAN: Okay.

THE COURT: And what he will then do is have the opportunity to point out to the jury what his mind is saying to him when he looks at it and why those are to him marks that are consistent marks, the same marks, consistent means the same. This mark here is consistent with the mark here, you can see it starts over here, runs all the way over there, it kind of gets scrambled here. This is a dark zone in here, we don't see it going over, but look down here we see it, all that is just comparison, known to unknown. And the jury can draw whatever inferences. And he gets the benefit of being able to explain -- I mean, he gets to talk about, he can talk about class characteristics and subclass characteristics and individual characteristics and the phraseology. He can talk about, you know, the microscope and how he got them and test fired, and

explain how he got these two things and how the comparison microscope works and the different powers and everything, and take the laser thing.

I don't know how we're going to get that precision on the screen with high def, but I leave that up to you. Maybe you'll have to bring in a HD TV in here or something, I don't know.

You can laser print and go through that, and the defense can move wherever it needs to move to see it, and he can spend whatever time is reasonable identifying how this mark connects with that mark, and we're looking over here, and you can see right here, and these, the marks are consistent. These are the marks that we know came from a cartridge fired from the gun that was recovered, and these are consistent with those, see how it goes there, there, there, that's what I'm allowing him to do.

MR. NAUVEL: Okay. And then, Your Honor, I honestly don't know whether this is required or not, but to the extent that the government needs to object to preserve any kind of --

THE COURT: Let me make it clear. My ruling is a Federal Rule of Evidence 103 pretrial definitive ruling on the record which preserves your ability to be able to say I'm completely wrong in the event that there's an appeal taken from this case by the government, and those who are smart and wiser than me down in Richmond can decide whether I made a mistake or

not. But by making it a pretrial definitive ruling on the record in accordance with Rule 103, it prevents you from having to offer the evidence, get an objection, and having me rule at trial.

MR. NAUVEL: Thank you, Your Honor.

THE COURT: And if you have any question about that, take a look at Rule 103 and you'll see it's not one of those, oh, I'm kind of leaning this way, but let's see what happens at trial. This is definitive in pretrial. The whole purpose of that change which came in in, I think 2000, or one of the 1990 changes was to preserve the record so that you now know it's a definitive ruling, not a "I'm inclined to" ruling.

MR. NAUVEL: Okay, thank you, Your Honor.

THE COURT: And it is important to preserve it, so I think you've done that.

Mr. Oppenheimer.

MR. OPPENHEIMER: I think you addressed it, Judge, the use of the word "identified" is what made me stand.

THE COURT: We've identified that as a word that's not going to be used.

MR. OPPENHEIMER: And then I guess I just want to be sure that the opinion is couched in terms of an opinion. The language we've been using is that the marks are consistent, insinuating are, in fact, consistent with as opposed to his opinion.

THE COURT: Well, I think that the way to queue it up is for counsel to say, do you have an opinion whether the marks that you've identified as having been made by the test fire of the gun are consistent with the marks that you described to the jury on the bullets recovered at the scene?

Yes, I have an opinion.

What is your opinion?

It's consistent.

Do you wish to explain?

And by then he would have gone through it all again. And he's not going to repeat it all, but that's the way to phrase it so that the question is a specific response to a question posed by counsel, and I think that should address your issue.

MR. OPPENHEIMER: I have nothing further.

THE COURT: So, on the Rule 61 was the defendant's prior convictions. We've agreed that we're not --

Mr. McVeigh, thank you very much, sir.

MR. McVEIGH: Thank you, Your Honor.

THE COURT: I appreciate, you had a long couple of days, and I appreciate your perseverance, sir.

MR. McVEIGH: Thank you.

THE COURT: We've agreed that the homicide conviction is not going to be brought in for 404(b) purposes, right?

MR. NAUVEL: Yes, Your Honor.

THE COURT: And we've agreed that we're not going to mention parole status.

MR. NAUVEL: Yes, Your Honor.

THE COURT: But we do have to get the felony conviction in for the felony in possession, and I thought that the only open issue was, are you working on a stipulation of that?

MR. NAUVEL: Yes, we're working on stipulations, and we apologize, Your Honor, we actually have not, all four of us had an opportunity --

THE COURT: I understand that.

MR. NAUVEL: But we do, having discussed with the defense, we do anticipate having these issues resolved before the time --

THE COURT: I just don't want to miss it between now and -- we have one more hearing that we have to get through.

MR. NAUVEL: Right.

THE COURT: And I am planning on writing an opinion on that one.

MR. NAUVEL: Okay.

THE COURT: Now, Ms. Kaplan, are you here just as a guest appearance or are you -- I understand Mr. Sullivan has been chased out to the Senate Judiciary Committee.

MS. KAPLAN: I am in this for the long haul, Your Honor, and I'm attempting to get up to speed as quickly as

possible.

THE COURT: Well, you have -- there's an old expression that you'd be blessed with an interesting case. This is an interesting case. You have very skillful, but very professional opposing counsel, and you have a very skillful and very articulate co-counsel, so I have no doubt that in the time that we have remaining that everyone will be able to continue in that vein with the excellent work that both sides have been doing.

Mr. Lassiter.

MR. LASSITER: I was only standing in regards to the stipulation question. I didn't know whether you wanted to hear from my side or not. I don't have anything to add at all.

THE COURT: You're working on the issue.

MR. LASSITER: Yes.

THE COURT: Okay. Anything further, folks? I think I've taken enough of your time the last couple of days.

Thank you all very much for your hard work. This is exhausting, and both of you have done fabulous -- all of you have done fabulous work on this, so thank you.

(Court recessed at 3:57 p.m.)

- o -

CERTIFICATE OF COURT REPORTER

I, Linda C. Marshall, certify that the foregoing is a correct transcript from the record of proceedings in the above-entitled matter.

/s/
_____
Linda C. Marshall, RPR
Official Court Reporter