# THE ADMISSIBILITY OF FIREARMS AND TOOLMARKS EXPERT TESTIMONY IN THE SHADOW OF PCAST

Colonel (Ret.) Jim Agar*



* Colonel James R. Agar II, U.S. Army (Retired). Colonel (Ret.) Agar currently serves as an Assistant General Counsel for the FBI's Forensic Laboratory at Quantico, Virginia. He also served over 25 years on active duty with the U.S. Army Judge Advocate General's Corps and is a Distinguished Graduate of the National War College in Washington, D.C. The opinions and comments contained in this article are solely those of the author. They do not reflect the official position of—nor are they endorsed by—the Department of Justice, the Federal Bureau of Investigation, or the FBI Forensic Laboratory.

## I. INTRODUCTION

Courts in the United States have routinely admitted expert testimony regarding firearms identification for over a century.[1] In the mid-1920s, modern techniques to examine and compare bullets and cartridge cases debuted, featuring trained forensic examiners who utilized a comparison microscope to conduct their examination.[2] This technique, pioneered by

[1] Commonwealth v. Best, 62 N.E. 748, 750 (Mass. 1902); State v. Clark, 196 P. 360, 367–69 (Or. 1921); Laney v. United States, 294 F. 412, 416 (D.C. Cir. 1923).

[2] 1 J. HOWARD MATHEWS, FIREARMS IDENTIFICATION, at xi (1962).

Calvin H. Goddard, underwent substantial scrutiny by trial judges and appellate courts in the late 1920s and early 1930s.[3] The courts found Goddard's methodology to be relevant and reliable, admitting his expert testimony and permitting him to testify that a specific bullet or cartridge case had been fired by a particular firearm.[4] He called the new discipline "Firearms Identification"[5] or "Forensic Ballistics."[6]

Goddard's methodology proved instrumental in solving some of the nation's biggest murder cases. Firearms identification figured prominently in the investigations of the St. Valentine's Day Massacre[7] and the murders committed by the Italian anarchists Sacco and Vanzetti.[8] Almost four decades after the Sacco and Vanzetti trial, Chief Justice Earl Warren relied on firearms examiners to determine whether a rifle owned by Lee Harvey Oswald fired the bullets that assassinated President John F. Kennedy.[9] For the next fifty-five years, U.S. courts admitted expert testimony in the area of firearms identification with experts who sometimes possessed little to no qualifications.

But in 2016, the little-known President's Council of Advisors on Science and Technology (PCAST) issued a report on forensic science that criticized the reliability of the firearms identification discipline, finding the discipline lacked "foundational validity."[10]

---

[3] *See* Evans v. Commonwealth, 19 S.W.2d 1091, 1093–99 (Ky. 1929); State v. Campbell, 239 N.W. 715, 719–25 (Iowa 1931).

[4] *Evans*, 19 S.W.2d at 1093–99; *Campbell*, 239 N.W. at 724.

[5] *Campbell*, 239 N.W. 715 at 719.

[6] 1 MATHEWS, *supra* note 2.

[7] WILLIAM J. HELMER & ARTHUR J. BILEK, THE ST. VALENTINE'S DAY MASSACRE 116, 158–60, 168, 180 (2004). Goddard's examinations excluded dozens of .45 caliber Thompson sub-machine guns over a period of months before identifying the specific firearm used to murder seven Chicago gangsters in 1929. *Id.* at 158, 180.

[8] FRANCIS RUSSELL, SACCO & VANZETTI: THE CASE RESOLVED 151–56, 158–62 (1986). Calvin Goddard's 1927 examination of the .32 pistol allegedly used in the murders and his conclusion that Sacco's pistol was the one used to commit the crime was the subject of controversy at the time. *Id.* 151–52. However, independent firearms examinations in 1961 and 1983 later confirmed Goddard's conclusion. *Id.* 158, 160.

[9] HON. EARL WARREN, C.J. ET AL., REPORT OF THE PRESIDENT'S COMMISSION ON THE ASSASSINATION OF PRESIDENT KENNEDY, 79–85, 547–62 (1964) [hereinafter WARREN COMMISSION REPORT], https://www.archives.gov/research/jfk/warren-commission-report.

[10] EXECUTIVE OFFICE OF THE PRESIDENT, PRESIDENT'S COUNCIL OF ADVISORS ON SCIENCE AND TECHNOLOGY,_REPORT TO THE PRESIDENT: FORENSIC SCIENCE IN CRIMINAL COURTS: ENSURING SCIENTIFIC VALIDITY OF FEATURE COMPARISON METHODS, 112 (2016) [hereinafter PCAST REPORT],

The PCAST Report followed two similar reports from the National Academy of Sciences in 2008[11] and 2009,[12] which also questioned the accuracy of the firearms identification discipline. The PCAST report has become the impetus for a small minority of courts to revisit the admissibility of the expert testimony of firearms examiners. In doing so, these courts are attempting to reverse nearly a century of jurisprudence by restricting or denying the ability of firearms expert witnesses to identify a specific firearm as the source of a fired bullet or cartridge case.[13] Does this constitute a new standard for the admissibility of firearms expert testimony, and do these courts provide sound reasoning?

This article analyzes the PCAST report on which those courts relied, examines the courts' rationale in restricting firearms experts' testimony, and questions whether such restrictions are warranted.

Before beginning that discussion, however, this article provides a general overview of the firearms identification discipline and its historic admissibility in the courts.

## II. FIREARMS EVIDENCE ANALYSIS

Contemporary firearms examinations closely follow the methodology Calvin Goddard pioneered nearly a century ago. During its investigation of President Kennedy's assassination, the Warren Commission described the fundamental principles of firearms identification as follows:

---

https://obamawhitehouse.archives.gov/sites/default/files/microsites/ostp/PCAST/pcast_forensic_science_report_final.pdf. (last visited June 24, 2020). PCAST also produced an Addendum to its report on January 6, 2017. *An Addendum to the PCSAT Report on Forensic Science in Criminal Court*s (Jan. 6, 2017), https://obamawhitehouse.archives.gov/sites/default/files/microsites/ostp/PCAST/pcast_forensics_addendum_finalv2.pdf.

[11] *See generally* NATIONAL RESEARCH COUNCIL, BALLISTIC IMAGING (2008) [hereinafter BALLISTIC IMAGING], https://www.nap.edu/download/12162.

[12] *See generally* HON. HARRY T. EDWARDS, ET AL. STRENGTHENING FORENSIC SCIENCE IN THE UNITED STATES: A PATH FORWARD (2009) [hereinafter NRC REPORT], https://www.ncjrs.gov/pdffiles1/nij/grants/228091.pdf.

[13] The cases which placed these restrictions on firearms expert witnesses are: Williams v. United States, 210 A.3d 734, 736, 738–43 (D.C. 2019); United States v. Davis, No. 4:18-cr-00011, 2019 U.S. Dist. LEXIS 155037, at *26 (W.D. Va. Sept. 11, 2019); United States v. Tibbs, No. 2016-CF1-19431, 2019 D.C. Super. LEXIS 9, at *23 (D.C. Sept. 5, 2019); United States v. Shipp, 422 F. Supp. 3d 762, 765–66 (E.D.N.Y. 2019); United States v. Adams, 444 F. Supp. 3d 1248 (D. Ore. 2020); and People v. Ross, 129 N.Y.S.3d 629 (N.Y. Sup. Ct. 2020).

> A cartridge, or round of ammunition, is composed of a primer, a cartridge case, powder, and a bullet. The primer, a metal cup containing a detonable mixture, fits into the base of the cartridge case, which is loaded with the powder. The bullet, which usually consists of lead or of a lead core encased in a higher strength metal jacket, fits into the neck of the cartridge case. To fire the bullet, the cartridge is placed in the chamber of a firearm, immediately behind the firearm's barrel. The base of the cartridge rests against a solid support called the breech face or, in the case of a bolt-operated weapon, the bolt face. When the trigger is pulled, a firing pin strikes a swift, hard blow into the primer, detonating the priming mixture. The flames from the resulting explosion ignite the powder, causing a rapid combustion whose force propels the bullet forward through the barrel.
>
> The barrels of modern firearms are "rifled," that is, several spiral grooves are cut into the barrel from end to end. The purpose of the rifling is to set the bullet spinning around its axis, giving it a stability in flight that it would otherwise lack. The weapons of a given make and model are alike in their rifling characteristics; that is, number of grooves, number of lands (the raised portion of the barrel between the grooves) and twist of the rifling. When a bullet is fired through a barrel, it is engraved with these rifling characteristics.
>
> In addition to rifling characteristics, every weapon bears distinctive microscopic characteristics on its components, including its barrel, firing pin, and breech face. While a weapon's rifling characteristics are common to all other weapons of its make and model (and sometimes even to weapons of a different make or model), a weapon's microscopic characteristics are distinctive, and differ from those of every other weapon, regardless of make and model. Such markings are initially caused during manufacture, since the action of manufacturing tools differs microscopically from weapon to weapon, and since the tools change microscopically while being operated. As a weapon is used,

> further distinctive microscopic markings are introduced by the effects of wear, fouling, and cleaning . . .
>
> . . . When a cartridge is fired, the microscopic characteristics of the weapon's barrel are engraved into the bullet (along with its rifling characteristics), and the microscopic characteristics of the firing pin and breech face are engraved into the base of the cartridge case. By virtue of these microscopic markings, an expert can frequently match a bullet or cartridge case to the weapon in which it was fired. To make such an identification, the expert compares the suspect bullet or cartridge case under a comparison microscope, side by side with a test bullet or cartridge case which has been fired in the weapon, to determine whether the pattern of the markings in the test and suspect items are sufficiently similar to show that they were fired in the same weapon.[14]

In conducting their examination, firearms examiners typically consider three different characteristics of the bullet or cartridge case.[15] These include class characteristics, individual characteristics, and the infrequently examined sub-class characteristics.[16]

Class characteristics describe a variety of distinctive, measurable, objective features, such as the caliber of the bullet or cartridge case, the material of the same, the firing pin impression, general rifling characteristics (five lands, left twist), breech-face marks, manufacturer identification, headstamp, bullet weight, and priming material.[17] These objective

[14] WARREN COMMISSION REPORT, *supra* note 9, at 547–48, 551. In the course of the investigation into the assassination of President Kennedy, one intact bullet was recovered from the stretcher bearing Texas Governor John Connelly, and five bullet fragments were recovered from the President's limousine. *Id.* at 79. Firearms identification experts were able to identify the intact bullet and two of the larger fragments as having been fired by the rifle found in the Texas School Book Depository. *Id.* at 85. Three cartridge cases found in the Depository were also identified as having been fired by the rifle found there. *Id.* at 79, 85.

[15] NRC REPORT, *supra* note 12, at 152.

[16] *Id.*

[17] BALLISTIC IMAGING, *supra* note 11, at 31–34, 46, 55, 58; FBI LABORATORY, FIREARMS/TOOLMARKS DISCIPLINE STANDARD OPERATING PROCEDURE FOR CARTRIDGE CASE EXAMINATIONS 1, 4 (Rev. 5, 2020) [hereinafter FBI CARTRIDGE SOP], https://fbilabqsd.com/; FBI LABORATORY, FIREARMS/TOOLMARKS UNIT STANDARD OPERATING PROCEDURE FOR BULLET EXAMINATIONS 1, 3 (Rev. 5, 2020) [hereinafter FBI BULLET SOP], https://fbilabqsd.com/.

characteristics result from design factors and are determined prior to the manufacturing of a firearm.[18] While class characteristics may be useful in eliminating a bullet or cartridge case as being fired from a particular firearm, or in restricting the pool of potential firearms which could have fired a bullet or cartridge case, firearms examiners cannot use these characteristics to identify a particular bullet or cartridge case's source.[19]

Individual characteristics are the marks considered unique to an individual tool or firearm.[20] These marks include either random imperfections or irregularities incidental to manufacturing or are caused by use, corrosion, or damage.[21]

"Sub-class" characteristics straddle the divide between class and individual characteristics.[22] Produced during the firearm's manufacturing, sub-class characteristics may be common to an extremely small group of firearms, such as those which are manufactured contemporaneously on the same assembly line.[23] An examiner might confuse sub-class characteristics, which more than one firearm may possess, with individual characteristics, which are unique to a particular firearm. Yet sub-class characteristics typically do not pose a problem for an examiner because they are

[18] FBI CARTRIDGE SOP, *supra* note 17, at 1, 4; FBI BULLET SOP, *supra* note 17, at 1, 3.

[19] BALLISTIC IMAGING, *supra* note 11, at 57–58. For example, an examiner receives a 9mm cartridge case recovered from the crime scene, but a .357 revolver was recovered at the suspect's house. *See id.* Because the class characteristics of the two are not in agreement, the .357 revolver can be eliminated as the source of the 9mm cartridge case. *See id.* Even if a 9mm pistol had been recovered from a suspect's house, the class characteristics cannot tell the examiner it was THAT 9mm pistol which fired the cartridge case. *See id.* More examination would be needed to ascertain whether the cartridge and pistol shared any individual characteristics before the examiner could entertain the possibility of identifying the recovered pistol as the source which fired the cartridge case. *See id.*

[20] NRC REPORT, *supra* note 12, at 152.

[21] FBI CARTRIDGE SOP, *supra* note 17, at 1; FBI BULLET SOP, *supra* note 17, at 1. Individual characteristics are marks produced by the random imperfections or irregularities of tool surfaces. FBI CARTRIDGE SOP, *supra* note 17, at 1; FBI BULLET SOP, *supra* note 17, at 1. These random imperfections or irregularities are produced incidental to manufacture and/or caused by use, corrosion, or damage. FBI CARTRIDGE SOP, *supra* note 17, at 1; FBI BULLET SOP, *supra* note 17, at 1. They are considered unique to that tool to the practical exclusion of all other tools. NRC REPORT, *supra* note 12, at 152.

[22] NRC REPORT, *supra* note 12, at 152.

[23] *Id.* Sub-class characteristics are also features that may be produced incidental to manufacturing and which are consistent among a small number of items fabricated by the same tool in the same approximate state of wear. *Id.* These features are not determined prior to manufacture and are more restrictive than class characteristics. *Id.*

dimensionally larger than individual characteristics and only occur over a short duration for a handful of firearms on an assembly line.[24]

Regardless, "[t]he task of the firearms and toolmark examiner is to identify the individual characteristics of microscopic toolmarks apart from class and subclass characteristics and then to assess the extent of agreement in individual characteristics in the two sets of toolmarks to permit the identification of an individual tool or firearm."[25]

The Federal Bureau of Investigation (FBI) permits firearms examiners to reach three conclusions or opinions.[26] First, a "Source Exclusion" allows an examiner to opine that the firearm examined could not have fired the bullets or cartridge case in question.[27] A "Source Exclusion" is defined as the examiner's opinion that two bullets or cartridge cases did not come from the same source or firearm.[28] Second, an examiner can identify the bullet or cartridge case as having been fired by a particular firearm, which is referred to as "Source Identification."[29] The Department of Justice (DOJ) defines "Source Identification" in the following manner:

> [A]n examiner's conclusion that two toolmarks originated from the same source. This conclusion is an examiner's opinion that all observed class characteristics are in agreement and the quality and quantity of corresponding individual characteristics is such that the examiner would not expect to find that same combination of individual characteristics repeated in another source and has found

[24] James E. Hamby et al., *The Identification of Bullets Fired from 10 Consecutively Rifled 9mm Ruger Pistol Barrels: A Research Project Involving 507 Participants from 20 Countries*, 41 No. 2 ASS'N FIREARM AND TOOL MARK EXAM'RS J., 99, 104, 107 (2009) [hereinafter *Hamby Study*].

[25] NRC REPORT, *supra* note 12, at 153. In addition, sub-class characteristics appear to have little, if any, impact in how firearms examiners reach their conclusions, or the accuracy of those conclusions. *See Hamby Study*, *supra* note 24.

[26] FBI LAB'Y, FIREARMS/TOOLMARKS DISCIPLINE, FBI APPROVED STANDARDS FOR SCIENTIFIC TESTIMONY AND REPORT LANGUAGE, 2–3 Sept. 20, (Rev. 4 2020) [hereinafter FBI ASSTR], https://fbilabqsd.fbi.gov/file-repository/firearms—toolmarks/quality-assurance/12-ftd-fbi-aprvd-stndrds-for-scientfc-testimony-and-rpt-language-firearms-toolmarks-discipline-4.pdf/view (last visited Aug. 7, 2021).

[27] *Id.* at 2.

[28] *Id.* at 2. *See also* DEP'T OF JUST., UNIFORM LANGUAGE FOR TESTIMONY AND REPORTS FOR THE FORENSIC FIREARMS/TOOLMARKS DISCIPLINE PATTERN EXAMINATION 2 (2020) [hereinafter DOJ FIREARMS ULTR], https://www.justice.gov/olp/uniform-language-testimony-and-reports.

[29] FBI ASSTR, *supra* note 26, at 2; DOJ FIREARMS ULTR, *supra* note 28, at 2.

> insufficient disagreement of individual characteristics to conclude they originated from different sources.
>
> The basis for a 'source identification' conclusion is an examiner's opinion that the observed class characteristics and corresponding individual characteristics provide extremely strong support for the proposition that the two toolmarks originated from the same source and extremely weak support for the proposition that the two toolmarks originated from different sources.[30]

It should also be noted that an examiner's "Source Identification" conclusion "is not based upon a statistically-derived or verified measurement or an actual comparison to all other firearms or toolmarks in the world."[31]

Third, an examiner may opine that his or her examination or comparison is "inconclusive," because while the observed class characteristics agree, there is insufficient quality and/or quantity of corresponding individual characteristics that the examiner is unable to identify or exclude the two tool marks as having originated from the same source.[32] "Reasons for an 'inconclusive' conclusion include the presence of microscopic similarity, . . . a lack of any observed microscopic similarity; or microscopic dissimilarity that is insufficient to form the conclusion of 'source exclusion.'"[33]

The definitions outlined in the DOJ's Firearms Uniform Language of Testimony and Reporting (ULTR) is not a "methodology" for conducting an examination. Instead, it describes uniform terms and definitions, plus the conceptual approach and bases for the conclusions drawn. It only regulates forensic examinations conducted by the DOJ.[34] The DOJ also imposed several limitations on examiners' reports and in-court testimony when rendering conclusions. For example, firearms and tool marks examiners cannot testify that their "source identification" opinion excludes all other firearms in the world.[35] Nor can the examiner declare that the firearms

[30] DOJ FIREARMS ULTR, *supra* note 28, at 2.

[31] *Id.* at 3.

[32] FBI ASSTR, *supra* note 26, at 3; *see also* DOJ FIREARMS ULTR, *supra* note 28, at 2.

[33] DOJ FIREARMS ULTR, *supra* note 28, at 3; *see also* FBI ASSTR, *supra* note 26, at 3.

[34] DOJ FIREARMS ULTRA, *supra* note 28, at 1. The DOJ Firearms ULTR is only binding on forensic firearms examiners in the Department of Justice, to include the FBI, ATF, and DEA. *Id.* State and local crime labs remain free to adopt their own limitations and reporting language. *Id.*

[35] *Id.* at 3.

identification discipline has a zero percent error rate or give a numerical weight to their "source identification" opinion or conclusion.[36]

By comparison, many firearms examiners at state and local crime labs utilize the Theory of Identification from the Association of Firearms and Toolmarks Examiners (AFTE) as a methodology by which they conduct their examinations. The AFTE Theory of Identification, which closely parallels the DOJ ULTR, "enables opinions of common origin to be made when the unique surface contours of two toolmarks are in 'sufficient agreement,'"[37] which exists between two tool marks when "the agreement of individual characteristics is of a quantity and quality that the likelihood another tool could have made the mark is so remote as to be considered a practical impossibility."[38] AFTE cautions that an identification is subjective and reflects the examiner's training and experience.[39]

## III. FIREARMS EXPERT TESTIMONY UNDER FRE 702; PRE-2009

During the first five decades of firearms identification expert testimony, its admissibility in court was governed by the 1923 case of *Frye v. United States*.[40] To admit the testimony of an expert witness, the *Frye* test required a scientific principle or discovery to be "sufficiently established to have gained general acceptance in the particular field in which it belongs."[41] Firearm identification experts encountered little difficulty in clearing the *Frye* threshold. Indeed, Calvin Goddard's early cases in the late 1920s and early 1930s served as a blueprint for other courts to evaluate the firearms identification discipline and the expertise of any witness called to testify in that field.[42]

[36] *Id.*

[37] *AFTE Theory of Identification as it Relates to Toolmarks*, ASS'N OF FIREARMS AND TOOLMARKS EXAM'RS, https://afte.org/about-us/what-is-afte/afte-theory-of-identification (last visited June 17, 2020).

[38] *Id.*

[39] *Id.*

[40] *See generally* 293 F. 1013 (App. D.C. 1923).

[41] *Id.* at 1014.

[42] *See* Evans v. Commonwealth, 19 S.W.2d 1091, 1093–99 (Ky. 1929); State v. Campbell, 239 N.W. 715, 719–25 (Iowa 1931).

In 1975, the Federal Rules of Evidence (the "Rules") took effect for all federal courts,[43] and many state courts adopted the Rules soon thereafter. The 1975 Rules included Rule 702, which specifically addressed the admissibility of expert opinion testimony.[44] Even after adopting Rule 702, courts continued to admit firearms identification evidence and found such evidence to be reliable.[45] The current text of Rule 702, titled "Testimony by Expert Witnesses," was last amended in 2011 and reads as follows:

> A witness who is qualified as an expert by knowledge, skill, experience, training, or education may testify in the form of an opinion or otherwise if:
>
> (a) the expert's scientific, technical, or other specialized knowledge will help the trier of fact to understand the evidence or to determine a fact in issue;
>
> (b) the testimony is based on sufficient facts or data;
>
> (c) the testimony is the product of reliable principles and methods; and
>
> (d) the expert has reliably applied the principles and methods to the facts of the case.[46]

The 2000 amendment to Rule 702 was in response to two landmark U.S. Supreme Court cases on expert testimony: *Daubert v. Merrell Dow Pharmaceuticals*[47] and *Kumho Tire Co. v. Carmichael*.[48] In *Daubert*, the district court granted a motion to exclude expert testimony concerning the

[43] FED. R. EVID. 702 (1975) (repealed 2000); Pub. L. No. 93-12, 87 Stat. 9 (1973); Pub. L. No. 93-595, 88 Stat. 1926 (1975). The Federal Rules of Evidence became effective on July 1, 1975. Pub. L. No. 93-595, 88 Stat. 1926 (1975).

[44] FED. R. EVID. 702 (1975) (repealed 2000).

[45] United States v. Bowers, 534 F.2d 186, 193 (9th Cir. 1976), *cert. denied*, 429 U.S. 942 (1976) ("The record was sufficient to permit the trial court to conclude that 'tool mark identification' rests upon a scientific basis and is a reliable and generally accepted procedure.").

[46] FED. R. EVID. 702. The original Rule 702 read as follows: "If scientific, technical, or other specialized knowledge will assist the trier of fact to understand the evidence or to determine a fact in issue, a witness qualified as an expert by knowledge, skill, experience, training, or education, may testify thereto in the form of an opinion or otherwise." FED. R. EVID. 702 (1975) (repealed 2000).

[47] 509 U.S. 579 (1993).

[48] 526 U.S. 137 (1999).

effects of a drug called Bendectin on children in utero.[49] The U.S. Court of Appeals for the Ninth Circuit affirmed the decision, finding the experts' methodology was not "generally accepted as reliable in the relevant scientific community"[50] and that the basis for the experts' opinion was "unpublished, not subjected to the normal peer review process and generated solely for use in litigation."[51]

In a unanimous opinion, the U.S. Supreme Court reversed the Ninth Circuit's decision.[52] In *Daubert*, the Court rejected the seventy-year-old *Frye* test as superseded by the Rules twenty years before.[53] The Court also found that federal judges have a "gatekeeping"[54] role under Rule 702 to ensure expert testimony is both relevant and reliable.[55] Courts have long dealt with the concept of relevance; reliability, however, was another matter. To assist the lower courts in gauging the reliability of a scientific theory or technique and evaluating the admissibility of expert testimony, the Court provided five non-binding factors.[56]

The first factor considers whether a scientific theory or technique "can be (and has been) tested."[57] The second factor asks "whether the theory or technique has been subjected to peer review and publication."[58] The third factor involves any "known or potential rate of error."[59] The fourth factor weighs "the existence and maintenance of standards controlling the technique's operation."[60] And finally, the fifth factor evaluates the "general acceptance" in the "relevant scientific community."[61] The Court cautioned that the "focus, of course, must be solely on principles and methodology, not on the conclusions that they generate,"[62] and it emphasized that the inquiry

[49] *Daubert*, 509 U.S. at 583–84.
[50] *Id.* at 584.
[51] *Id.*
[52] *Id.* at 598.
[53] *Id.* at 587–89.
[54] *Id.* at 597.
[55] *Id.* at 589–91.
[56] *Id.* at 592–94.
[57] *Id.* at 593.
[58] *Id.* at 593–94.
[59] *Id.* at 594.
[60] *Id.*
[61] *Id.*
[62] *Id.* at 595.

to determine reliability under Rule 702 is a "flexible one."[63] The Court also noted its decision was limited to scientific testimony, not testimony based on "technical or other specialized knowledge."[64] Finally, the Justices held that concerns over a more relaxed standard of admitting expert witness testimony were overly pessimistic about the capabilities of the jury and of the adversary system generally.[65] Vigorous cross examination, presentation of contrary evidence, and careful instruction on the burden of proof are the traditional and appropriate means of attacking shaky but admissible evidence.[66] The *Daubert* factors soon took on a life of their own, and consequently, the lower courts sometimes misread the application of both Rule 702 and the *Daubert* decision. In 1997, the Supreme Court revisited *Daubert* in *General Electric v. Joiner*.[67] Here the Court announced that, just like every other evidentiary ruling at trial, an appellate court would review a district court's decision to admit or exclude expert witness testimony under an abuse-of-discretion standard.[68] The Court further held that "the Federal Rules of Evidence allow district courts to admit a somewhat broader range of scientific testimony than would have been admissible under *Frye*."[69]

The Court revisited *Daubert* two years later in *Kumho Tire Co. v. Carmichael*.[70] The *Kumho* Court specifically addressed testimony that was based on technical or specialized knowledge but was not scientific.[71] In *Kumho Tire*, the Court considered whether the district court abused its discretion by excluding a tire failure expert's testimony as to the cause of the plaintiff's automobile accident.[72] The Eleventh Circuit ruled it had.[73] The U.S. Supreme Court, however, reversed the Eleventh Circuit's decision, holding that *Daubert*'s "gatekeeping" obligation "applies not only to testimony based on 'scientific' knowledge, but also to testimony based on 'technical' and 'other specialized' knowledge . . . .[A] trial court *may*

[63] *Id.* at 594.

[64] *Id.* at 590, n.8.

[65] *Id.* at 595–96.

[66] *Id.* at 596.

[67] General Electric v. Joiner, 522 U.S. 136, 138–39 (1997).

[68] *Id.* at 143, 146.

[69] *Id.* at 142.

[70] Kumho Tire Co. v. Carmichael, 526 U.S. 137, 140 (1999).

[71] *Id.* at 142, 146–47.

[72] *Id.* at 142–46.

[73] Carmichael v. Samyang Tire, Inc., 131 F.3d 1433 (11th Cir. 1997), *rev'd sub nom.* Kumho Tire Co. v. Carmichael, 526 U.S. 137 (1999).

consider one or more of the more specific factors that *Daubert* mentioned when doing so will help determine that testimony's reliability."[74]

The Court also held that, irrespective of the *Daubert* factors, "the relevant reliability concerns may focus upon personal knowledge or experience."[75] Shortly thereafter, Rule 702 was amended in 2000 to reflect the Court's decisions in *Daubert* and *Kumho Tire*.[76] With ample guidance from the U.S. Supreme Court, federal and state courts soon began using the metric of Rule 702 to determine the reliability and admissibility of firearms identification experts' testimony.[77]

In 2004, the U.S. Court of Appeals for the Fifth Circuit held in *United States v. Hicks* that the district court did not abuse its discretion by admitting the expert witness testimony of a firearms identification examiner under Rule 702.[78] The appellate court also found that "the matching of spent shell casings to the weapon that fired them has been a recognized method of ballistics testing in this circuit for decades."[79] The court assessed the *Daubert* factors in reaching its decision, examining the error rates, published firearms studies, and the general acceptance of the firearms identification discipline.[80] In 2007, the U.S. Court of Appeals for the Second Circuit in *United States v. Williams* also upheld the district court's decision to admit the testimony of a firearms identification expert witness, despite the lack of a formal pre-trial admissibility hearing.[81] Citing *Daubert*, the appellate court found the admission of the firearm expert's testimony was primarily based on her

[74] *Kumho Tire*, 526 U.S. at 141.

[75] *Id.* at 150.

[76] FED. R. EVID. 702 (amended 2000).

[77] United States v. Hicks, 389 F. 3d 514, 526 (5th Cir. 2004); United States v. Santiago, 199 F. Supp. 2d 101, 110–12 (S.D.N.Y. 2002). Ironically, the U.S. Supreme Court in *United States v. Scheffer* had already implied—in dicta—that DNA, fingerprints, and ballistics were reliable, setting them apart from polygraph evidence because these are "experts witnesses who testify about factual matters outside the jurors' knowledge . . . ." 523 U.S. 303, 313 (1998).

[78] 389 F. 3d 514, 526 (5th Cir. 2004).

[79] *Id.*

[80] Of some concern is the examiner's statement in this case that there was a "zero" or "near zero" rate of error for the discipline of firearms identification. *Id.* at 526. No discipline is capable of that degree of accuracy, and the U.S. Supreme Court recognized as much in the *Daubert* case: "it would be unreasonable to conclude that the subject of scientific testimony must be 'known' to a certainty; arguably, there are no certainties in science." *Id.* at 590.

[81] 506 F.3d 151, 161–62 (2nd Cir. 2007).

education, training, and experience,[82] which was sufficient to satisfy the judge's gatekeeping inquiry under Rule 702.[83]

A survey of reported opinions from U.S. district courts and state courts from 2000–2008 reveals many of the courts reviewed the admissibility of firearms identification expert testimony. One of these early cases was *United States v. Santiago*,[84] where the Southern District of New York opined expert testimony for firearms identification would be admissible even if such expertise was not from the "scientific community"[85] and "was based purely on experience."[86] No pre-trial admissibility hearing was held in *Santiago*.[87] Yet the trial court relied, in part, on the implicit endorsement of firearms expert witnesses by the U.S. Supreme Court in *United States v. Scheffer*,[88] where the Court upheld the exclusion of polygraph evidence at a court-martial because a polygraph examiner was "unlike other expert witnesses who testify about factual matters outside the jurors' knowledge, such as the analysis of fingerprints, ballistics, or DNA found at a crime scene . . . "[89]

In *United States v. Foster*, the U.S. District Court for the District of Maryland held a pre-trial admissibility hearing, yet the court found the testimony of an FBI firearms expert to be admissible without restrictions.[90] The same held true in *United States v. Natson*, where the U.S. District Court for the Middle District of Georgia admitted the expert testimony of another FBI firearms examiner without restrictions.[91]

However, some courts began to question the admissibility of firearms identification expert testimony. In *United States v. Green*, the district court allowed the firearms expert witness to testify with the limitation that any

[82] *Id.* at 161.

[83] *Id.* at 161–62.

[84] 199 F. Supp. 2d 101, 110–12 (S.D.N.Y. 2002). It should be noted that *Santiago* was not the first reported U.S. District Court case to consider an admissibility challenge to firearms identification evidence. That distinction probably goes to *United States v. Cooper*, 91 F. Supp. 2d 79, 82–84 (D.D.C. 2000), where the district court denied the defense challenge to the admissibility of firearms identification expert testimony.

[85] *Santiago*, 199 F. Supp. 2d at 112.

[86] *Id.*

[87] *See id.* at 110.

[88] *Id.* at 112.

[89] United States v. Scheffer, 523 U.S. 303, 313 (1998).

[90] 300 F. Supp. 2d 375, 376–77 (D. Md. 2004) (The court also determined that such opinion testimony was available even when the expert witness did not have a known firearm available for examination or comparison.).

[91] 469 F. Supp. 2d 1253, 1261 (M.D. Ga. 2007).

"match" found between the questioned cartridge cases and the known cartridge cases were not "to the exclusion of all other guns."[92] The court's chief concern hinged on problems with the firearms examiner's credibility.[93] The examiner took no notes, photos, or drawings during his examination.[94] He cited no error rates, professional certification, or proficiency testing regarding his ability as an examiner.[95] The examiner held to no specific protocol in conducting his examinations, and the laboratory where he performed the examination was neither certified nor accredited by any independent organization.[96] These issues factored into the court's decision to limit the expert's testimony.[97]

Perhaps the most exhaustive admissibility analysis of firearms expert testimony occurred in *United States v. Monteiro*, where the U.S. District Court for the District of Massachusetts held a six-day admissibility hearing on the issue of firearms identification.[98] At the conclusion of this hearing, the court found "the underlying scientific principle that firearms leave unique marks on ammunition is reliable under Rule 702."[99] The court also found firearms expert testimony reliable because of the publication and peer review of the AFTE Journal, an acceptably low error rate,[100] and general acceptance in the relevant scientific community.[101] However, the judge expressed concerns about the AFTE Theory of Identification because it was "tautological."[102] As in *United States v. Green*,[103] the court identified multiple problems with the firearms examiner because he had no formal academic or scientific training, lacked any professional certification, was not a member of any professional organization, reviewed no literature in his field, and did not take—let alone pass—a proficiency test in firearms identification at the time he conducted his examination.[104] Worse, the examiner failed to

[92] 405 F. Supp. 2d 104, 124 (D. Mass. 2005).

[93] *See id.* at 120–21.

[94] *Id.* at 113.

[95] *Id.* at 109, 116.

[96] *Id.* at 109–10.

[97] *Id.* at 120–22.

[98] 407 F. Supp. 2d 351, 355 (D. Mass. 2006).

[99] *Id.* at 366.

[100] *Id.* at 366–68.

[101] *Id.* at 371–72.

[102] *Id.* at 370.

[103] 405 F. Supp. 2d 104 (D. Mass. 2005).

[104] *Monteiro*, 470 F. Supp. 2d at 373.

document his analysis with sketches or photos as required by the AFTE; additionally, the examiner failed to have a second independent examiner review or verify his analysis.[105] The court ruled the examiner's testimony was inadmissible under Rule 702.[106] The court found the methodology of firearms identification reliable; however, the examiner's failure to adequately document the bases for his conclusions in his report in accordance with AFTE protocols made verification of his conclusions by another expert virtually impossible.[107] The court permitted the prosecution to supplement the record so the examiner could bring his examination up to "established standards in the field."[108] If he were able to do so, then he could testify that the recovered cartridge cases came from a particular firearm, with a caveat that he testify to a "reasonable degree of ballistic certainty."[109] Both *Green* and *Monteiro* were primarily concerned with the reliability (or the lack thereof) of the *examiner's credentials* and analysis of the evidence.[110] Neither of these decisions imposed substantive or material changes to the testimony of the examiners, nor did it alter their conclusions which identified a particular firearm as having fired a specific bullet or cartridge case.

Later, the U.S. District Court for the Northern District of California in *United States v. Diaz* found that "no reported decision has ever excluded firearms-identification expert testimony under *Daubert*."[111] While the court admitted the testimony of the firearms expert and allowed the examiner to

[105] *Id.* at 374.

[106] *Id.*

[107] *See id.* at 368–69, 374.

[108] *Id.* at 375.

[109] *Id.* (The court held specifically "Defendants' motion in limine to exclude ballistics evidence is ALLOWED without prejudice to supplementation by the government . . . The government must ensure that its proffered firearms identification testimony comports with the established standards in the field for peer review and documentation. If the expert opinion meets these standards, the expert may testify that the cartridge cases were fired from a particular firearm to a reasonable degree of ballistic certainty. However, the expert may not testify that there is a match to an exact statistical certainty."). Today, this testimony would not be allowed by the Department of Justice, which prohibits testifying to a "reasonable [degree of] scientific certainty" or words to that effect. *See* Memorandum from Loretta Lynch, Att'y Gen., to Heads of Dep't Components (Sept. 6. 2016) [hereinafter Lynch Memo], https://www.justice.gov/opa/file/891366/download, last visited June 24, 2020.

[110] *See generally Monteiro*, 407 F. Supp. 2d 351; United States v. Green, 405 F. Supp. 2d 104 (D. Mass. 2005) (emphasis added).

[111] United States v. Diaz, No. CR 05-00167 WHA, 2007 U.S. Dist. LEXIS 13152, *14 (N.D. Cal. Feb. 12, 2007).

opine as to the origin of a bullet or cartridge case, the examiner could not testify that his identification conclusion was "to the exclusion of all other firearms in the world."[112]

The 2008 case *United States v. Glynn* was the first case to impose any substantive change to the opinion offered by a firearms examiner.[113] In *Glynn*, the U.S. district court judge found "that whatever else ballistics identification analysis could be called, it could not fairly be called 'science.'"[114] The court analyzed the firearm identification discipline under *Kumho Tire*, wherein the judge acknowledged "its methodology has garnered sufficient empirical support as to warrant its admissibility."[115] The court then restricted the examiner in the case from testifying that a bullet or cartridge case "matched" or that it came from a particular firearm, permitting the examiner only to testify that any match was "more likely than not."[116]

This decision is troubling for three reasons. First, the judge was doing far more than imposing a limitation on the expert witness. In fact, the court *rewrote* the substantive testimony of a witness, essentially scripting what the witness would testify to under oath.[117] It mattered not that the witness's opinion differed with what the judge scripted or that the witness believed he identified the source of the bullet or cartridge case.[118] The judge forced the witness to testify differently.[119] Confronted with the phrase "more likely than not," what if a witness were to respond, "But, your honor, that is not my opinion. It is yours."?[120]

Second, no scientific foundation supports the term "more likely than not." The court in *Glynn* cited no scientific standard or firearms study which evaluated or supported a conclusion of "more likely than not."[121] Third, much has been written about the phrase "to a reasonable degree of scientific certainty" and how such a phrase cannot be used to quantify the weight to be given an expert's opinion. Yet "more likely than not" is equally misleading. The jury has no idea what to make of such testimony. Is it fifty percent more

[112] *Id.* at 35–36, 41–42.

[113] *See generally* 578 F. Supp. 2d 567 (S.D.N.Y. 2008).

[114] *Id.* at 570.

[115] *Id.* at 574.

[116] *Id.* at 574–75.

[117] *See id.*

[118] *See id.*

[119] *See id.*

[120] *See generally id.*

[121] *See id.*

likely, seventy-nine percent, or even ninety-nine percent? The jury has no basis for determining what weight to give the expert's testimony. And while the witness has testified "more likely than not," that was *not* the expert's opinion. It was the opinion of the court that the witness was required to parrot back to the jury.

Third, the *Glynn* decision stands as an anomaly in its own circuit, contrary to established precedent from the U.S. Court of Appeals for the Second Circuit. The appellate court had previously affirmed the unrestricted admissibility of firearms identification expert testimony—just a year before *Glynn* was decided—in *United States v. Williams*.[122] And the same appellate court would do so again a decade later in *United States v. Gil*.[123]

The *Glynn* case marks a significant departure from over seventy-five years of established jurisprudence regarding firearms identification expert testimony. For the first time in American history, a judge ordered an expert witness in the area of firearms identification to make substantive and material changes to their expert opinion and directed the expert to not identify the source of a cartridge case or bullet—contrary to the conclusions in the witness's written report.

## IV. THE NATIONAL RESEARCH COUNCIL REPORTS (2008 & 2009)

In 2008, just before the court decided *Glynn*, the National Academy of Sciences National Research Council (NRC) published the first of two studies addressing the issue of firearms identification.[124] In BALLISTIC IMAGING, the NRC commissioned a review to assess the feasibility, accuracy, and technical capability of a national ballistics database to aid criminal investigations.[125] The DOJ's National Institute of Justice (NIJ) sponsored the report with the support of the National Institute of Standards and Technology (NIST) from the Department of Commerce.[126] The committee conducting the review included Lawden Yates, a former firearms examiner, forensic laboratory director, and general counsel to the Alabama Department of Forensic Sciences.[127] Assisting the committee were the Bureau of Alcohol, Tobacco, & Firearms (ATF); the President of the Association of Firearms and

[122] 506 F.3d 151, 161–62 (2nd Cir. 2007).

[123] No. 16-524, 2017 WL 689719, at *1 (2d Cir. 2017).

[124] BALLISTIC IMAGING, *supra* note 11.

[125] *Id.* at 2.

[126] *Id.* at xi.

[127] *Id.* at xii.

Toolmarks Examiners (AFTE); and scores of firearms examiners from state and federal forensic laboratories.[128]

The NRC Committee ultimately concluded that a national ballistics image database was not feasible at that time.[129] In doing so, it also found that the "validity of the fundamental assumptions of uniqueness and reproducibility of firearms-related toolmarks has not yet been fully demonstrated."[130] The BALLISTIC IMAGING report went on to declare that "firearms-related toolmarks are not completely random and volatile; one can find similar marks on bullets and cartridge cases from the same gun."[131] The NRC also took pains to note that "this study is neither a verdict on the uniqueness of firearms-related toolmarks generally nor an assessment of the validity of firearms identification as a discipline."[132] In fact, the BALLISTIC IMAGING report was limited to automated imaging systems with the potential to create a national database.[133] It did not study the abilities of trained human examiners.[134] Furthermore, "the proposal for this study explicitly precluded the committee from assessing the admissibility of forensic firearms evidence in court, either generally or in specific regard to testimony on ballistic imaging comparisons."[135]

---

[128] *Id.* at xi–xiv.

[129] *Id.* at 5.

[130] *Id.* at 3. The report also states:

> We also note that the committee does not provide an overall assessment of firearms identification as a discipline nor does it advise on the admissibility of firearms-related toolmarks evidence in legal proceedings: these topics are not within its charge. The committee's charge is to determine the extent to which the toolmarks left on bullets and cartridge casings after firing a weapon can be captured by imaging technology. It is also to assess whether a ballistic image database—particularly a national RBID containing images of exhibits fired from all newly manufactured and imported guns—would be feasible and operationally useful, by which we mean capable of generating leads for follow-up and further investigation.

*Id.* at 3–4. "[T]he proposal for this study explicitly precluded the committee from assessing the admissibility of forensic firearms evidence in court, either generally or in specific regard to testimony on ballistic imaging comparisons." *Id.* at 20.

[131] *Id.* at 3.

[132] *Id.* at 18.

[133] *See id.* (for a general statement of the report's purpose).

[134] *Id.* at 19.

[135] *Id.* at 20. Prof. John Rolph of the University of Southern California chaired the NRC Committee which conducted the analysis and authored the BALLISTIC IMAGING report. *See* Affidavit of Dr. John E. Rolph, United States v. Edwards, No. F-516-01 (D.C. Super. Ct. 2008),

A year later, in 2009, the NRC published a study of forensic science titled STRENGTHENING FORENSIC SCIENCE IN THE UNITED STATES: A PATH FORWARD.[136] This congressionally mandated 286-page report reviewed a broad spectrum of forensic science, with a particular focus on firearms identification and several other disciplines.[137] The committee authoring this report included no less than four people with firsthand experience in the field of forensic science.[138] It also heard from dozens of presenters, representing virtually every discipline in forensic science.[139] The committee submitted thirteen recommendations to Congress and the Department of Justice regarding ways to improve forensic science.[140] None of the recommendations specifically mentioned firearms identification.[141] Yet the report itself criticized firearms and toolmarks, not to mention almost every other forensic discipline: "With the exception of nuclear DNA analysis, however, no forensic method has been rigorously shown to have the capacity to consistently, and with a high degree of certainty, demonstrate a connection between evidence and a specific individual or source."[142]

The report questioned whether courts were sufficiently equipped to evaluate some forensic evidence, particularly firearms and toolmarks expert testimony, using the case of *United States v. Green*[143] as an example.[144] The report addressed the admissibility of firearms expert testimony and concluded: "we must limit the risk of having the reliability of certain forensic

https://afte.org/uploads/documents/swggun-rolph-affidavit.pdf. He declared in an affidavit that the NRC's BALLISTIC IMAGING report should not be construed as a comment on the admissibility nor reliability of firearms identification. *See id.*

[136] NRC REPORT, *supra* note 12.

[137] *Id.* at 3 (Some of the other disciplines reviewed by the NRC REPORT included: Friction ridge (fingerprint) analysis, biological evidence (DNA), impression evidence (shoeprint and tire tread analysis), bitemarks, questioned document examination, and hair & fiber analysis.).

[138] *Id.* at 293, 296–97 (Retired FBI employee, Prof. Randall Murch worked at the FBI Forensic Laboratory as a forensic examiner. Prof. Jay Siegel is a former forensic chemist and a member of both the International Association for Identification and an Academic Affiliate member of the American Society of Crime Lab Directors. He also published two textbooks on forensic evidence. Peter M. Marone was the director of the Virginia Department of Forensic Sciences. Prof. Robert Shaler is a professor of forensic science at Penn State University.).

[139] *Id.* at xi–xii.

[140] *Id.* at 19–33.

[141] *Id.*

[142] *Id.* at 7.

[143] 405 F. Supp. 2d 104 (D. Mass. 2005).

[144] NRC REPORT, *supra* note 12, at 107–08.

science methodologies [condoned by the courts] before the techniques have been properly studied and their accuracy verified."[145]

The NRC REPORT then turned its attention to the firearms and toolmarks discipline. It noted that "even with more training and experience using newer techniques, the decision of the toolmark examiner remains a subjective decision based on unarticulated standards and no statistical foundation for estimation of error rates."[146] The report also took issue with the lack of studies in the firearms and toolmarks discipline and asserted that additional studies in this field were warranted to assess its reliability.[147] This finding nested perfectly with the NRC REPORT's third recommendation that additional published, peer-reviewed studies be funded to develop "quantifiable measures of the reliability and accuracy of forensic analyses."[148] Finally, the NRC REPORT criticized the lack of a precisely defined protocol or process in the field of firearms identification and the ambiguity of the term "sufficient agreement" used in the AFTE Theory of Identification.[149]

[145] *Id.* at 12.

[146] *Id.* at 153–54.

[147] *Id.* at 154 ("Toolmark and firearms analysis suffers from the same limitations discussed above for impression evidence. 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 the 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 wear 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.").

[148] *Id.* at 22–23.

[149] *Id.* at 155 ("A fundamental problem with toolmark and firearms analysis is the lack of a precisely defined process. As noted above, AFTE has adopted a theory of identification, but it does not provide a specific protocol. It says that an examiner may offer an opinion that a specific tool or firearm was the source of a specific set of toolmarks or a bullet striation pattern when 'sufficient agreement' exists in the pattern of two sets of marks. It defines agreement as 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.' The meaning of 'exceeds the best agreement' and 'consistent with' are not specified, and the examiner is expected to draw on his or her own experience. This AFTE document, which is the best guidance available for the field of toolmark identification, does not even consider, let alone address, questions regarding variability, reliability, repeatability, or the number of correlations needed to achieve a given degree of confidence.").

## V. ADMISSIBILITY OF FIREARMS EXPERT TESTIMONY DURING THE INTER-REPORT YEARS (2008–2016)

Following the publication of the NRC's reports on ballistic imaging and forensic science, some courts more closely scrutinized the firearms identification discipline. In *United States v. Willock*, the U.S. District Court for the District of Maryland considered the two NRC reports in reaching its decision.[150] The court allowed the firearms examiner to opine that the examined firearm was the source of the cartridge case at the crime scene; however, the court prohibited the expert witness from testifying that it was a "practical impossibility" for any other firearm to have fired the cartridge case, nor could the expert give any degree of certainty as to his opinion.[151] The court recommended these limitations because the examiner who testified at trial had not personally examined all the cartridge cases in evidence to reach his conclusion that there was a match.[152] Instead, he utilized the observations of another examiner as the foundation for his expert witness opinion as to firearm identification.[153]

Other federal district courts required similar limitations. In *United States v. Ashburn*, the court found firearms identification expert testimony passed all the *Daubert* factors and was admissible under Rule 702.[154] Nevertheless, it limited the expert's testimony by not permitting him to say he was "100% certain" or that it was a "practical impossibility" that another firearm could have fired the recovered items, or that his identification was to "the exclusion of all other firearms in the world."[155] Though the court, like many others, permitted him to testify that his conclusions were to "a reasonable degree of

[150] F. Supp. 2d 536, 555–56 (D. Md. 2010). In this case the U.S. District Court judge adopted the recommendation of Magistrate Judge Paul Grimm. *Id.* at 549–82.

[151] *Id.* at 549.

[152] *Id.* at 546, 573–74.

[153] *Id.* 573–74 ("[I]t is impossible to appreciate that Sgt. Ensor and his colleagues did not personally compare all the cartridges in Case No. 1496 against all of the cartridges in Case No. 93. Beyond the limited physical examination he did make of the evidence in the City Case, Sgt. Ensor assumes that the cartridges in Case No. 1496 matched the City cartridges he did not examine, because a Baltimore City toolmark examiner—whose qualifications, proficiency, and adherence to proper methodology is unknown—said they did . . . Accordingly, because Sgt. Ensor's own opinion, for lack of a more elegant expression, piggybacks on those of an unknown Baltimore City examiner, I recommend that Sgt. Ensor not be able to express his opinions to the same degree of certainty as other courts have permitted."). *See also Willock*, 682 F. Supp. 2d at 535–36 (where the district court judge accepted the magistrate's recommendation regarding limitations on the expert's testimony).

[154] United States v. Ashburn, 88 F. Supp. 3d 239, 247 (E.D.N.Y. 2015).

[155] *Id.* at 249.

ballistics certainty."[156] Similarly, in *United States v. Taylor*,[157] the court permitted the expert to opine as to the source of the bullet but precluded him from testifying that it was the source "to the exclusion . . . of all other" firearms and that his conclusion was to a "reasonable degree of certainty in the firearms examination field."[158]

Yet in the 2013 case of *United States v. Stafford*, the U.S. Court of Appeals for the Sixth Circuit held that the firearms expert's testimony at trial as to gunshot residue and firearm identification was admissible under Rule 702.[159] And in *United States v. Wrensford*, the U.S. District Court for the Virgin Islands also admitted the expert testimony of a firearms identification expert witness without restriction.[160]

In 2015, the U.S. Court of Appeals for the Ninth Circuit addressed the issue of firearm identification expert testimony in *United States v. Cazares*.[161] The court approved the utilization of toolmark identification testimony by requiring the expert to testify that the conclusions were to "a reasonable degree of certainty in the ballistics field."[162] A year after the Ninth Circuit's decision, the DOJ rejected this characterization, and in 2016, Attorney General Loretta Lynch ordered forensic examiners to refrain from using the terms "reasonable scientific certainty," "reasonable [degree of firearms discipline] certainty," or words to that effect.[163]

---

[156] *Id.* at 250.

[157] 663 F. Supp. 2d 1170 (D.N.M. 2009). Only one person has been deemed as not reliable as an expert witness under FRE 702 on the subject of firearms identification, thereby precluding her testimony before a jury on the topic. That person is Prof. Adina Schwartz of the John Jay College of Criminal Justice. In *United States v. Taylor*, a U.S. District Court judge in New Mexico refused to recognize Prof. Schwartz as an expert witness because "She has no experience in conducting firearms or toolmark identification examinations, nor has she ever taken a proficiency test in the field of firearm investigations; indeed she testified before this Court that she has never fired a gun." 704 F. Supp. 2d 1192 at 1195–96, 1199–1200 (D.N.M. 2009). Other courts have found Prof. Schwartz is not a neutral scholar on the subject of firearms identification evidence, but an advocate against its admissibility. *See United States v. Otero*, 849 F. Supp. 2d 425, 437 (D.N.J. 2012), *aff'd*, 557 F. App'x 146 (3d Cir. 2014) ("Professor Schwartz's opinions are substantially outweighed by the evidence supporting admissibility.").

[158] *Taylor*, 663 F. Supp. at 1180.

[159] 721 F.3d 380, 392–95 (6th Cir. 2013).

[160] No. 2013-0003, 2014 U.S. Dist. LEXIS 102446, at *60 (D.V.I. July 28, 2014).

[161] 788 F.3d 956, 989 (9th Cir. 2015).

[162] *Id.*

[163] *See* Lynch Memo, *supra* note 109. In addition, the DOJ's ULTR pertaining to the firearms and toolmarks disciplines also proscribes the use of the term "to a reasonable degree of scientific (or forensic) certainty." *See* DOJ FIREARMS ULTR, *supra* note 28, at 3.

Other federal district courts refused to impose any limitations on firearms identification expert testimony. For example, in *United States v. Casey*, the U.S. District Court for the District of Puerto Rico denied a defense request to exclude firearms identification expert testimony.[164] In the District of New Jersey, the court in *United States v. Otero* denied the defense's request to exclude firearms identification evidence, finding the discipline's methodology sound and the testimony admissible under Rule 702.[165] Practically ignoring the decision in *Glynn*[166] from another federal judge in the same jurisdiction four years before, the court placed no restrictions on the expert's opinion testimony.[167] The *Otero* court's decision is all the more striking because, unlike the judge's ruling in *Glynn*, the *Otero* court had the benefit of reviewing both the NRC's 2008 *Ballistics Imaging* and 2009 *Strengthening Forensic Science in the United States: A Path Forward* (*Strengthening Forensic Science*) reports along with the testimony of a defense expert witness in the case.[168]

During this eight-year period, a number of state courts also considered the implications of the two NRC reports and the critical rulings in the *Glynn* case. All ruled that firearms identification expert testimony was admissible with few, if any, limitations. This included courts in Alabama,[169] Arizona,[170]

[164] 928 F. Supp. 2d 397, 399–400 (D.P.R. 2013).

[165] 849 F. Supp. 2d 425, 437–38 (D.N.J. 2012), *aff'd*, 557 F. App'x 146 (3d Cir. 2014).

[166] 578 F. Supp. 2d 567, 574–75 (S.D.N.Y. 2008).

[167] *See Otero*, 849 F. Supp. 2d at 438.

[168] *Id.* at 430.

[169] Revis v. State, 101 So. 3d 247, 290–92 (Ala. Crim. App. 2011).

[170] State v. Romero, 341 P.3d 493, 499 (Ariz. Ct. App. 2014); State v. Foshay, 370 P.3d 618, 622–24 (Ariz. Ct. App. 2016).

Maryland,[171] New York,[172] Ohio,[173] Indiana,[174] Massachusetts,[175] Illinois,[176] and the District of Columbia.[177] Notedly, the D.C. Court of Appeals later performed a partial retreat on the admissibility of firearms identification testimony in the 2016 case of *Gardner v. United States*, where the court held it was error for the trial court to admit the "unqualified" expert opinion of a firearms identification expert witness or testify to "100% certainty" that an expert witness identified a particular firearm as the source of cartridge case or bullet.[178]

## VI. FIREARMS STUDIES AFTER THE NRC REPORTS: 2009–2016

Meanwhile, the firearms identification discipline began an intense period of reflection during the seven-year interval between the release of the 2009 NRC Report and the 2016 release of the PCAST Report,[179] culminating in the

[171] Patterson v. State, 146 A.3d 496, 501–04 (Md. Ct. Spec. App. 2016); Fleming v. State, 1 A.3d 572, 589–91 (Md. Ct. Spec. App. Md. 2010). The reader should note that, until recently, Maryland did not follow *Daubert* or FRE 702. It adhered to the *Frye* standard for admissibility of expert witness testimony. *See* Reed v. State, 391 A.2d 364, 368 (Md. 1978), *overruled by* Rochkind v. Stevenson, 236 A.3d 630 (Md. 2020).

[172] People v. Givens, 912 N.Y.S.2d 855, 857 (Sup. Ct. 2010). New York continues to follow the *Frye* standard and follows neither *Daubert* nor FRE 702. *See* People v. Powell, No. 22, 2021 WL 5407448, at *5 (N.Y. Nov. 18, 2021).

[173] State v. Langlois, 2 N.E.3d 936, 950–51 (Ohio Ct. App. 2013). The appellate court conducted a detailed analysis of the law, the NRC reports, and the discipline of firearms identification before reaching its decision. *See id.* at 944–47. "Our conclusion on this issue finds support in the decisions of other appellate districts in Ohio, notwithstanding the recent criticisms in scientific reports and the limitations some federal courts have imposed on the testimony of firearms experts. These decisions hold that the methodology of comparatively analyzing and testing bullets and shell cases recovered from crime scenes is reliable." *Id.* at 950.

[174] Turner v. State, 953 N.E.2d 1039, 1053–54 (Ind. 2011).

[175] Commonwealth v. Pytou Heang, 942 N.E.2d 927, 946–47 (Mass. 2011).

[176] People v. Robinson, 2 N.E.3d 383, 402 (Ill. App. Ct. 2013).

[177] Jones v. United States, 27 A.3d 1130, 1136–38 (D.C. 2011). At the time of the *Jones* decision, D.C. was a *Frye* jurisdiction. It abandoned *Frye* in 2016 and embraced both *Daubert* and FRE 702 in the case of Motorola Inc. v. Murray, 147 A.3d 751 (D.C. 2016).

[178] Gardner v. United States, 140 A.3d 1172, 1184 (D.C. 2016) ("[W]e now hold that the trial court erred by allowing Mr. Watkins to give an unqualified opinion about the source of the bullet that killed Mr. Kamara. We further hold that in this jurisdiction a firearms and toolmark expert may not give an unqualified opinion, or testify with absolute or 100% certainty, that based on ballistics pattern comparison matching a fatal shot was fired from one firearm, to the exclusion of all other firearms.").

[179] PCAST REPORT, *supra* note 10.

publication of six new studies addressing the reliability of firearms identification. The studies responded to the NRC's two reports: the 2008 *Ballistic Imaging* and the 2009 *Strengthening Forensic Science* reports. These reports recommended that more study was needed to assess the reliability of firearms identification analysis, reporting, and testimony.

Studies of firearms identification were nothing new. The first recorded examination and comparison of cartridge cases fired from multiple firearms occurred in the United States at the U.S. Army's Frankford Arsenal in 1907.[180] Later, Calvin Goddard conducted one of the earliest studies of firearms identification in 1926.[181] From then until 2009, forensic scientists performed no less than forty-three validation studies of the firearms and toolmarks identification discipline.[182] Despite the wealth of firearms studies before 2009, firearms examiners commissioned the next generation of firearms identification studies to answer the central questions posed by the NRC reports: could a trained firearms and toolmark examiner reliably identify a bullet or cartridge case fired by a specific firearm?

This simple question raises complex issues. With over 310 million firearms in the U.S. alone,[183] no firearms examiner can possibly state they have compared the suspect cartridge case or bullet to every other firearm in existence or declare an identification "to the exclusion of every other firearm." Nor can any study review all other firearms in the world. Yet studies can utilize firearms equipped with barrels and slides that are consecutively manufactured on an assembly line and are mechanically identical. Consecutively manufactured barrels and slides represent the best possibility for the production of two firearms that could produce virtually indistinguishable markings on fired bullets and cartridge cases because the same tools and machining processes are utilized back-to-back on one barrel

[180] *Hamby Study*, *supra* note 24, at 100.

[181] *See* 1 MATHEWS, *supra* note 2, at 3. Goddard's 1926 study at Springfield Arsenal consisted of four consecutively manufactured gun barrels, with rounds fired through each barrel for comparison. *Id.* He found that each barrel produced bullets that did not match those produced by the other barrels, demonstrating the individual characteristics particular to each barrel. *Id.* He then fired 500 rounds from the barrel of a machine gun and found the first bullet could be matched to bullet number 500, demonstrating the persistence of the barrel's individual characteristics. *Id.*

[182] *Hamby Study*, *supra* note 24, at 100–04. Why none of these forty-three validation studies in firearms and toolmarks were mentioned in either of the reports from the NRC remains puzzling.

[183] WILLIAM J. KROUSE, CONG. RSCH. SERV., RL32842, GUN CONTROL LEGISLATION 8 (2012).

(or slide) after another.[184] "If there were ever any chance for duplication of individual marks, it would occur during the manufacture of consecutively manufactured barrels."[185]

Guns made with consecutively manufactured barrels or slides should mark every bullet and cartridge case virtually the same as the other barrels and slides from that same assembly line. Under these circumstances, if firearms examiners can discriminate between bullets and cartridge cases fired from guns with consecutively manufactured barrels and slides, then the hypothesis of firearm identification has been validated.

### *A. The Hamby Study (2009)*

Dr. James Hamby, David Brundage, and Dr. Jim Thorpe published the first of these studies in 2009.[186] The *Hamby Study* was a continuation of a ten-year study where 502 qualified firearms examiners each reviewed fifteen bullets fired through a Ruger P85 9mm semi-automatic pistol. In addition, the sample bullets introduced an element of difficulty not typically seen in casework. The study utilized not just ten different barrels for the Ruger, but ten *consecutively manufactured* barrels, with each barrel represented in the sample of fifteen bullets furnished to the examiners.[187] The study tested the hypothesis of whether firearms examiners could accurately determine which of the fifteen unknown or "questioned" bullets they examined were a match to the twenty "known" bullets in the study, each of which was fired from one of the ten consecutively manufactured barrels. The design of this study required examiners to control for possible sub-class characteristics. Yet of the 7,605 bullets examined by the 502 examiners, none of the participants committed an error, and all correctly linked each unknown bullet to the known bullet fired from the same barrel. Thus, there were no "false positives."[188] In 2019, Hamby published an update to his study in the *Journal*

[184] THOMAS G. FADUL, JR. ET AL., AN EMPIRICAL STUDY TO IMPROVE THE SCIENTIFIC FOUNDATION OF FORENSIC FIREARM AND TOOL MARK IDENTIFICATION UTILIZING CONSECUTIVELY MANUFACTURED GLOCK EBIS BARRELS WITH THE SAME EBIS PATTERN 37 (2013), https://www.ojp.gov/pdffiles1/nij/grants/244232.pdf [hereinafter *Miami-Dade Study*].

[185] *Id.*

[186] *Hamby Study*, *supra* note 24, at 99. This study was peer reviewed. *Id.*

[187] *Id.* at 104–05. Some barrels had as many as three bullets in the sample; however, each of the ten consecutively manufactured barrels had at least one bullet in the sample given to examiners. *Id.* at 105.

[188] *Id.* at 107. Eight of the bullets' examinations were deemed inconclusive because the examiner could neither identify nor exclude the bullet as having been fired by a particular barrel,

*of Forensic Sciences*.[189] It revealed a new total of 10,455 examinations by the 697 firearm examiners who participated in the study, all of whom successfully completed the test.[190] Once again, none of the participants reported a single missed identification or a "false positive" result.[191]

### *B. The First Fadul Study (2013)*

In 2013, Dr. Thomas Fadul and his colleagues published another study, this time using cartridge cases fired through ten consecutively manufactured 9mm Ruger slides.[192] This *Fadul Study* sought to determine whether qualified examiners, who had completed a training program at an accredited laboratory, could correctly identify which cartridge case came from one of ten consecutively manufactured slides. While attempting to falsify the AFTE Theory of Identification and the concept of "sufficient agreement," the *Fadul Study* recruited 217 firearm examiner participants, who made a total of 3,255 comparisons of cartridge cases (fifteen each). They produced 3,239 correct answers, two incorrect answers, and fourteen inconclusive determinations.[193] This translated into a misidentification error rate of 0.000636, or less than .07%.[194]

Fadul acknowledged that this study, like that of the 2009 *Hamby Study*, was a "closed set" in which the answer was always present in the fifteen cartridge cases submitted for examination to each test taker. While the participants were not told the examination was a closed set, they could deduce

reducing the number of correct identifications from 7,605 to 7,597. *Id.* By using consecutively manufactured barrels, Hamby introduced subclass characteristics as a factor to potentially confuse or deceive the examiners; however, subclass proved to be virtually no issue at all. *Id.*

[189] James E. Hamby et. al., *A Worldwide Study of Bullets Fired from 10 Consecutively Rifled 9mm RUGER Pistol Barrels—Analysis of Examiner Error Rate*, 64 J. FORENSIC SCIS. 551 (2019).

[190] *Id.* at 556.

[191] *Id.*

[192] Thomas G. Fadul, Jr. et al., *An Empirical Study to Improve the Scientific Foundation of Forensic Firearm and Tool Mark Identification Utilizing 10 Consecutively Manufactured Slides*, 45 AFTE J., 376–93 (2013) [hereinafter *Fadul Study*]. The "slide" is the top part of a semi-automatic pistol which houses the breech face, firing pin, and ejector pin, which typically create the toolmarks on a cartridge case when it is ejected from the firearm. Fadul's study was commissioned by the DOJ's National Institute of Justice. *Id.* This study was also peer reviewed. *See Peer Review Process*, ASS'N OF FIREARM & TOOL MARK EXAM'RS, https://afte.org/afte-journal/afte-journal-peer-review-process.

[193] *Id.* at 386.

[194] *Id.* at 384–85. Fadul defines an "error rate" as "a calculated value that represents the comparison of the number of wrong responses with the total number of responses." *Id.* at 382.

it from the answers as they matched each unknown cartridge case to a known cartridge case.[195] Fadul recommended the use of an "open set" design study in future research on firearm identification.[196] Despite the use of a closed set, both the Hamby and Fadul studies utilized either consecutively manufactured barrels or slides. The effect of this study design was such that both individual and sub-class characteristics were present in the studies, forcing the examiner to discriminate between the two and rely on the individual characteristics present on the bullets or cartridge cases they examined.

### *C. The Miami-Dade Study (2013)*

Fadul conducted a second similar study in 2013. This *Miami-Dade Study* reviewed the ability of firearms identification examiners to compare bullets fired through a special Glock barrel, featuring the Enhanced Barrel Identification System (EBIS) etched into the barrel by Glock.[197] The study's parameters required each of the 165 participating examiners to compare eight "known" pairs of bullets from each of the first eight EBIS barrels (sixteen "known" bullets) with ten "unknown" bullets, one fired through each of the ten consecutively manufactured EBIS barrels.[198] They were then asked to compare the two sets.[199] Unlike previous studies, this one constituted an "open set" design because the examiners would be unable to correctly identify two of the "unknown" bullets to any of the sixteen "known" bullets.[200] The answer was not in the set for those two rounds.[201] The study revealed a total of 1650 unknown fired bullets examined by the participants with "1,496 correct answers, 12 incorrect answers and 142 inconclusive

[195] *Id.* at 383.

[196] *Id.* at 388. Fadul specified future research should "[u]se an 'open set' design where the participant has no expectation that all questioned toolmarks should match one or more of the unknowns." *Id.* at 370. In other words, the correct answer may *not* be present in the materials submitted for examination.

[197] *Miami-Dade Study*, *supra* note 184. The study is referred to as the *Miami-Dade Study* because of the support the Miami-Dade police department provided, their experience in comparing Glock barrels, and the reviewed barrel being called the "Miami barrel."

[198] *Id.* at 28–30.

[199] *Id.* at 28–29.

[200] *Id.* at 28.

[201] *Id.* at 28–30.

answers."[202] Based on this data, Fadul calculated an error rate of 0.7%.[203] He did not include inconclusive answers in calculating the error rate as they were not considered errors.[204] The *Miami-Dade Study* revealed 142 inconclusive determinations by the examiners out of 1650 comparisons of the unknown fired bullets examined by the participants. This calculated to an inconclusive rate of 8.60%.[205]

### *D. The Stroman Study (2014)*

Angela Stroman, a criminalist with the California Department of Justice, Bureau of Forensic Services, published the next study a year later.[206] In this smaller-firearms identification study, Stroman employed three Smith and Wesson model 4006TSW semi-automatic pistols.[207] Twenty-five participants reviewed three "known" and three "unknown" cartridge cases fired from the three pistols.[208] In seventy-four out of seventy-five instances, the examiners correctly identified which firearm expelled the cartridge case.[209] There was one inconclusive result, yet there was virtually no error rate.[210] While this study design was a closed set and did not use consecutively manufactured slides, it demonstrated the consistent accuracy of firearms identification once again.

---

[202] *Id.* at 35.

[203] *Id.* at 33. Fadul also noted "[t]hese identifications are not absolute because it will never be possible to examine every firearm or tool in the world, a prerequisite to making absolute determinations." *Id.* at 34. The 1.2% error rate reported by Fadul is for the upper end of the 95% confidence interval. The point estimate for the error rate from the experimental data was 0.7%. *Id.* at 33.

[204] *Id.* at 32 ("[I]nconclusive responses are neither incorrect nor correct and may indeed be the most appropriate response in a situation in which the sample, lab policy, and/or examiner capabilities do not permit a more definitive conclusion.").

[205] *Id.* at 35. The inconclusive rate is determined by the total number of comparisons (142) which reached an opinion of "inconclusive" over the number of total reported comparisons (1,650).

[206] Angela Stroman, *Empirically Determined Frequency of Error in Cartridge Case Examinations Using a Declared Double-Blind Format*, 46 AFTE J. 157, 157–75 (2014) [hereinafter *Stroman Study*]. The Stroman Study was peer reviewed. *Id.* at 157.

[207] *Id.* at 162.

[208] *Id.* at 169.

[209] *Id.*

[210] *Id.*

### *E. The Baldwin (Ames) Study (2014)*

David Baldwin from the U.S. Department of Energy Laboratory in Ames, Iowa, created perhaps the most cited firearms study in recent history. The 2014 *Baldwin (Ames) Study* used twenty-five new Ruger SR9 semi-automatic pistols, with 218 firearms examiners participating in the study. [211] The examiners were provided with fifteen separate comparisons, consisting of one "unknown" and three "known" cartridge cases, which may or may not have been discharged by the source of the "unknown" cartridge case.[212] This was an "open set" design study in which the examiner could not use the process of elimination or deductive reasoning to eliminate or identify a particular cartridge case. The "known" and "unknown" cartridge cases with the same source appeared together in a set 1,090 times.[213] When that happened, the examiners incorrectly eliminated the source of the "unknown" cartridge case just four times, making for an observed false-negative rate of 0.367%.[214] When the four eliminations and the eleven inconclusive results were factored in, the examiners failed to identify the source 1.376% of the time.[215]

Cartridge cases with different sources (where the "known" and "unknown" cartridge cases were fired from different firearms) appeared in 2,180 other instances.[216] When this occurred, the examiners incorrectly called them "identifications" twenty-two times, but correctly determined them to be "eliminations" 1,421 times.[217] The other 735 determinations were listed as "inconclusive."[218] This yields an observed false-positive error rate of 1.01%. For technical reasons, the authors adjusted this value to report a maximum likelihood estimate of the average examiner false positive rate as 0.94%.[219]

[211] DAVID P. BALDWIN ET AL., AMES LAB., U.S. DEP'T OF ENERGY, A STUDY OF FALSE-POSITIVE AND FALSE-NEGATIVE ERROR RATES IN CARTRIDGE CASE COMPARISONS 3 (2014), https://apps.dtic.mil/dtic/tr/fulltext/u2/a611807.pdf [hereinafter *Baldwin Study*].

[212] *Id.* at 10–11.

[213] *Id.* at 15.

[214] *Id.*

[215] *Id.*

[216] *Id.* at 16.

[217] *Id.*

[218] *Id.*

[219] *Id.* at 16–17. The estimated 95% likelihood confidence bound interval for the error rate is 0.36% to 2.26%. *Id.* at 16.

They also provide an estimate for the "inconclusive" rate of 20.41%.[220] Inconclusive results were not counted as errors. There is no indication the *Baldwin (Ames) Study* was published or subjected to peer review.

### *F. The Smith Study (2016)*

Finally, a 2016 study conducted by Tasha Smith from the San Francisco Police Department Crime Lab[221] evaluated thirty-one firearms examiners in what is perhaps the most realistic representation of casework in any firearms identification study.[222] The *Smith Study* used eight .40 caliber semi-automatic pistols, with two made by each of Smith & Wesson, Glock, Sig Sauer, and Taurus.[223] Test fires from these weapons produced bullets and cartridge cases that were randomly mixed for each test kit.[224] Each test kit contained twelve bullets and twelve cartridge cases, with none of the test fires from consecutively manufactured barrels.[225] Due to the random nature of their selection, no two test kits were alike. The thirty-one examiners correctly identified 191 cartridge cases as coming from the source firearm, with just one false positive result.[226] They also correctly identified all 156 bullets to the source that fired them. This translated to a false-positive error rate of 0.14% for cartridge-case identifications and a 0.0% false positive rate for bullet identifications.[227] The *Smith Study* reported 204 "inconclusive" results;[228] however, the author noted the following:

> As indicated through this study, a conclusion of neither identification nor elimination adds weight and value to the clear response of identification or elimination. Examiners are trained to be more conservative when making their

[220] The inconclusive rate is determined by the total number of comparisons (735) that reached a determination of "exclusion" over the number of total reported comparisons (3,600).

[221] Tasha P. Smith et al., *A Validation Study of Bullet and Cartridge Case Comparisons Using Samples Representative of Actual Casework*, 61 J. FORENSIC SCIS. 939, 939–46 (2016) [hereinafter *Smith Study*].

[222] *Id.* at 940.

[223] *Id.* at 940–41.

[224] *Id.* at 941.

[225] *See id.* at 940.

[226] *Id.* at 943.

[227] *Id.* As for eliminations, the thirty-one examiners correctly eliminated 406 cartridge cases and 519 bullets with four false eliminations, for an error rate of 0.433% or less. *Id.*

[228] *Id.*

> evaluations and a response of inconclusive means that a particular examiner has not seen enough information to say that two items have been marked by the same tool or that they have not been marked by the same tool.[229]

Collectively, the sextet of firearms identification studies from 2009 to 2016 answered the call of the two NRC reports on firearms identification, confirming it as a reliable discipline with a consistently low error rate. Any time a firearms examiner made a conclusion or rendered an opinion of "identification" or opined a firearm was a "match" to a cartridge case or bullet, the examiner was accurate ninety-nine percent of the time.[230] The variously defined error rates for "identification" decisions by firearms examiners in these six studies can be summarized as follows:

*Hamby Study* =0%

*Fadul Study* =0.064%

*Stroman Study* = 0%

*Miami-Dade Study* = 0.7%

*Baldwin (Ames) Study* =1.01%

*Smith Study* = 0.14%

## VII. THE PCAST REPORT, 2016

Just six years after the publication of the congressionally mandated NRC report on forensic science, the PCAST undertook a critique of a number of "feature-comparison" disciplines in forensics, including latent fingerprints, DNA, shoe and tire tread analysis, and firearms and tool-marks identification.[231]

President Barack Obama commissioned PCAST in 2009 and issued an executive order to that effect a year later in 2010.[232] PCAST, "an advisory

[229] *Id.* at 945.

[230] *Hamby Study*, *supra* note 24, at 107; *Fadul Study*, *supra* note 192, at 384–85; *Stroman Study*, *supra* note 206, at 169; *Miami-Dade Study*, *supra* note 184, at 33; *Baldwin Study*, *supra* note 211, at 15–16; *Smith Study*, *supra* note 221, at 940–44.

[231] PCAST REPORT, *supra* note 10.

[232] Exec. Order No. 13,539, 75 Fed. Reg. 21,973 (Apr. 21, 2010). While the official website for PCAST lists 2009, this was actually the year the President announced the formation of PCAST

group of the nation's leading scientists and engineers who directly advise the President and the Executive Office of the President," makes "policy recommendations in the many areas where understanding of science, technology, and innovation is key to strengthening our economy and forming policy that works for the American people."[233] During its seven-year existence, PCAST issued thirty-nine reports on a wide array of topics, including cybersecurity, biological weapons, nanotechnology, spectrum policy, climate change, energy technologies, advanced manufacturing, ecosystems and the economy, antibiotic resistance, drug discovery and development, semiconductors, big data and privacy, pandemic flu vaccines, health information technology, STEM education, agriculture, and hearing aids.[234]

Yet PCAST's 2016 survey of forensic science created the most controversy. The 160-page report was titled "Forensic Science in Criminal Courts: Ensuring Scientific Validity of Feature Comparison Methods," and it evaluated seven feature-comparison disciplines to assess their scientific validity, including firearms and toolmark identification. [235]

PCAST first coined the new term "foundational validity" as a means to evaluate the reliability of a particular discipline.[236] As defined by PCAST, "foundational validity" consists of several requirements:

> *Foundational validity* for a forensic-science method requires that it be shown, based on empirical studies, to be *repeatable*, *reproducible*, and *accurate*, at levels that have been measured and are appropriate to the intended application. Foundational validity, then, means that a method can, *in principle*, be reliable. It is the *scientific*

under his administration. *See id.* The Executive Order, which created PCAST that would issue the 2016 report on forensic science, was E.O. 13539. *Id.*

[233] *About PCAST*, OBAMA WHITE HOUSE ARCHIVES, https://obamawhitehouse.archives.gov/administration/eop/ostp/pcast/about (last visited Jan. 15, 2020).

[234] *PCAST Documents & Reports*, OBAMA WHITE HOUSE ARCHIVES, https://obamawhitehouse.archives.gov/administration/eop/ostp/pcast/docsreports (last visited Jan. 15, 2020).

[235] PCAST REPORT, *supra* note 10, at 67–123. The six other "feature-comparison" disciplines PCAST evaluated in their report included: bitemarks analysis, latent fingerprint analysis, footwear analysis, hair analysis, DNA analysis of single source mixture samples, and DNA analysis of complex-mixture samples.

[236] *Id.* at 4–5.

> concept we mean to correspond to the *legal* requirement, in Rule 702(c), of "reliable principles and methods."[237]

The report then turned its attention specifically to firearms identification and rendered the following verdict: "firearms analysis currently falls short of the criteria for foundational validity, because there is only a single appropriately designed study to measure validity and estimate reliability. The scientific criteria for foundational validity require more than one such study, to demonstrate reproducibility."[238]

Simply put, the PCAST report determined that firearms identification expert testimony lacked the scientific foundation to be admissible. The report also criticized the firearms identification discipline for several other reasons. It found the discipline highly subjective and governed by the *AFTE Theory of Identification*, which PCAST considered to be circular reasoning.[239] The report dismissed many firearms studies as suffering from defects, such as closed set[240] or "white box" designs, which PCAST deemed not appropriate in assessing foundational validity or reliability.[241] Consequently, PCAST concluded that many of the firearms studies seriously underestimated the false-positive rate for firearms identification.[242] The report called for more studies in firearms identification, using appropriately designed "black box" studies, similar to the *Baldwin (Ames) Study*, and to convert firearms analysis from a subjective method to an objective one by using advances in technology.[243] Finally, the report recommended the following:

> If firearms analysis is allowed in court, the scientific criteria for validity as applied should be understood to require clearly reporting the error rates seen in appropriately designed black-box studies (estimated at 1 in 66, with a 95 percent confidence limit of 1 in 46, in the one such study to date).[244]

---

[237] *Id.*

[238] *Id.* at 112.

[239] *Id.* at 104.

[240] *Id.* at 109.

[241] *Id.* at 111.

[242] *Id.*

[243] *Id.* at 113.

[244] *Id.* at 150. It should be noted that the *Baldwin (Ames) Study* actually reported a false positive rate of only 94%, not the 2.2% false positive rate reported by PCAST. *Baldwin Study*, *supra* note 211, at 17.

## VIII. THE PROBLEMS WITH THE PCAST REPORT

While the PCAST Report examined several forensic disciplines, this paper focuses primarily on the review of the firearms and toolmarks discipline conducted by PCAST. To that end, the PCAST Report contains multiple problems that undermine the integrity of the report, rendering it an unreliable source—as a matter of science and law—to evaluate the firearms and toolmark discipline. These shortcomings include the makeup of persons who were affiliated with the PCAST Report, the use of terms and definitions alien to the firearms examination discipline or forensic science in general, and the use of arbitrary criteria to weigh the reliability of firearms analysis.

### *A. Lack of Qualifications by the Persons in PCAST*

The first of these problems is the composition of PCAST itself. While the Council included some members with knowledge in topics such as nanotechnology or big data, they lacked expertise in the field of firearms and toolmarks. When PCAST published the 2016 report on Feature-Comparison Methods, the Council consisted of two co-chairs, two vice-chairs, fifteen members, and a support staff of three. The PCAST working group for this report included five of the Council's members, plus one additional person who prepared the report, with the assistance of two additional staff members and a writer. A group of fourteen senior advisors—all judges and lawyers—assisted with the report.[245] In total, thirty-eight people researched, analyzed, drafted, and reviewed the PCAST Report on forensic science.[246]

Yet a review of the biographies of the thirty-eight individuals who assisted in the drafting and editing of the PCAST Report reveals the following:

[245] PCAST REPORT, *supra* note 10, at viii–ix. While the PCAST report on forensic science counted fourteen attorneys and judges among its senior advisors, other reports issued by PCAST had no attorneys advising the Commission in any capacity. These included PCAST's reports on safe drinking water, big data and privacy, and cybersecurity, to name a few. EXEC. OFF. OF THE PRESIDENT, PRESIDENT'S COUNCIL OF ADVISORS ON SCI. & TECH., REPORT TO THE PRESIDENT: IMMEDIATE OPPORTUNITIES FOR STRENGTHENING THE NATION'S CYBERSECURITY i–ii (2013); EXEC. OFF. OF THE PRESIDENT, PRESIDENT'S COUNCIL OF ADVISORS ON SCI. & TECH., REPORT TO THE PRESIDENT: SCIENCE AND TECHNOLOGY TO ENSURE THE SAFETY OF THE NATION'S DRINKING WATER v–viii (2016); EXEC. OFF. OF THE PRESIDENT, PRESIDENT'S COUNCIL OF ADVISORS ON SCI. & TECH., REPORT TO THE PRESIDENT: BIG DATA AND PRIVACY: A TECHNOLOGICAL PERSPECTIVE i–iii (2014).

[246] PCAST REPORT, *supra* note 10, at v–ix.

> None of these forty persons had any education, experience, or training in firearms identification or firearms evidence examination and analysis;
>
> None had ever worked for a firearms or ammunition manufacturer;
>
> None had been or were employed at a forensics laboratory;
>
> None had ever been a director of a forensics laboratory;
>
> Only two claimed to have any background in forensics whatsoever;[247]
>
> None had ever taken a proficiency test or competency test in firearms identification;
>
> None were certified as firearms identification examiners by the AFTE;
>
> None were AFTE members;
>
> None had been published in the area of firearms examination analysis;
>
> None possessed a degree in mechanical engineering, materials science, or metallurgy;[248]
>
> None had ever testified as an expert witness in firearms analysis or identification, and none had been recognized by

[247] *Id.* at 23 n.17; *PCAST Members*, OBAMA WHITE HOUSE ARCHIVES, https://obamawhitehouse.archives.gov/administration/eop/ostp/pcast/about/members (last visited Jan. 23, 2020). PCAST co-chair Eric Lander testified twice on DNA. PCAST REPORT, *supra* note 10, at 23. He and S. James Gates also served as members of the now-defunct National Commission on Forensic Science. *Id.* Neither cited any experience in the field of firearms and toolmark identification.

[248] In some cases, firearms experts have been qualified with merely a background in metallurgy or mechanical engineering and without any specialized knowledge of firearms. In Lee v. Smith & Wesson Corp., 760 F.3d 523, 525–26 (6th Cir. 2014), the appellate court found that a mechanical engineer was qualified to testify as an expert in firearms, albeit in the scope of a product liability case. In Seamon v. Remington Arms Co., 813 F.3d 983 (11th Cir. 2016), the court overturned a district court's decision to deny the admissibility of the expert testimony of a metallurgist on the cause of a malfunctioning firearm. *See* Seamon v. Remington Arms Co., 51 F. Supp. 3d 1198, 1204 (M.D. Ala. 2014). This, too, was a product liability case and not a case of firearms identification.

> any court as an expert witness in the area of either firearms analysis or firearms identification;
>
> None had conducted a firearms examination or authored a case report;
>
> None had conducted any empirical studies in the area of firearms analysis or identification;
>
> None had stated they had ever handled or fired a firearm;
>
> At least fourteen were judges or lawyers.

Notably, not a single person from the DOJ, law enforcement, or the ranks of state and local prosecutors were enlisted to draft or edit the PCAST Report.[249] The total absence of anyone with a background in forensic firearms and tool mark examinations among PCAST's thirty-eight members and staff renders the Council's review of that discipline highly suspect.[250] If a trial court judge were to assess the competency of PCAST's thirty-eight members and staff under Federal Rule of Evidence 702, it would likely find *none* of them competent as experts and thus would not permit any to testify before a jury on the topic of firearms identification evidence. How then could any of these thirty-eight people be competent to determine the "foundational validity" of the firearms and toolmark discipline? The question is particularly relevant when one considers that this report claims to be based on science, not legal analysis. Despite the presence of twenty lawyers and judges on the PCAST staff, the report took no position on the admissibility of firearms identification expert testimony at trial, stating, "[w]hether firearms analysis should be deemed admissible based on current evidence is a decision that belongs to the courts."[251]

### *B. Use of the Term "Metrology" as Applied to Firearms Examinations*

Because of PCAST's institutional lack of familiarity with the firearms discipline, cracks in its analysis of firearms and toolmark identification

[249] *See* PCAST REPORT, *supra* note 10, at v–ix, 23 n.17; *PCAST Members*, OBAMA WHITE HOUSE ARCHIVES, *supra* note 247.

[250] The reader will recall that both the NRC's 2008 *Ballistic Imaging* and 2009 *Strengthening Forensic Science* reports had firearms experts as part of the staff reviewing the data and writing the reports.

[251] PCAST REPORT, *supra* note 10, at 112.

quickly appeared. Initially, the PCAST Report claimed firearms identification (and other feature-comparison disciplines) belonged to the field of "metrology."[252] "Metrology" is defined as "the science that deals with measurement."[253] Yet the PCAST Report cites no authority for its determination that firearms analysis belongs to the field of metrology. While the NRC's 2008 *Ballistic Imaging Report* mentions "metrology" six times, the 2009 *Strengthening Forensic Science Report* never uses the word "metrology."[254] The AFTE does use the word "metrology" in some of its materials, but it is only in the context of algorithms which one day might be used to measure and compare toolmarks at some time in the future.[255] And while the PCAST Report referred to six studies of firearms identification, none of those six studies used the word "metrology." The only attempt by PCAST to cite a source or reference tying "metrology" to firearms identification is in a footnote of the PCAST Report itself, where PCAST claimed the National Institute of Standards and Technology (NIST) is researching forensic science. [256] However, the websites cited by PCAST never use the word "metrology" and only refer to measurements as being a *possible* means of identification of firearms in the future.[257] Accordingly, it appears the PCAST Report concluded by fiat that firearms identification is currently governed by "metrology." PCAST does not cite any scientific or legal source to support this conclusion.

The DOJ soundly rejected PCAST's failed attempt to bootstrap firearms examinations into the field of "metrology." It found the following:

---

[252] *Id.* at 23, 44.

[253] *Metrology*, THE AMERICAN HERITAGE DICTIONARY OF THE ENGLISH LANGUAGE (6th ed. 2016).

[254] BALLISTIC IMAGING, *supra* note 11; NRC REPORT, *supra* note 12.

[255] *See* ASS'N FIREARM & TOOL MARK EXAM'RS, https://afte.org/ (last visited Jan. 14, 2022).

[256] PCAST REPORT, *supra* note 10, at 44 n.93. The two NIST websites referenced by the PCAST Report can be found at https://www.nist.gov/about-nist/our-organization/mission-vision-values and https://www.nist.gov/topics/forensic-science. *NIST Mission, Vision, Core Competencies, and Core Values*, NAT'L INST. STANDARDS & TECH. (last visited Jan. 29, 2020); *Forensic Science*, NAT'L INST. STANDARDS & TECH. (last visited Jan. 29, 2020). More importantly, the NIST website for ballistics also lacks any use of the term "metrology" in its description of firearms identification. *See Ballistics*, NAT'L INST. STANDARDS & TECH., https://www.nist.gov/topics/ballistics (last visited Jan. 29, 2020).

[257] *How Good a Match is It? Putting Statistics into Forensic Firearms Identification*, NAT'L INST. STANDARDS & TECH. (Feb. 8, 2018), https://www.nist.gov/news-events/news/2018/02/how-good-match-it-putting-statistics-forensic-firearms-identification.

> Traditional forensic pattern examination methods—as currently practiced—do *not* belong to the scientific discipline of metrology. Forensic examiners visually *compare* the individual features observed in two examined samples, they do not *measure* [them.] The result of this comparison is a conclusion that is stated in words (nominal terms), not magnitudes (measurements).[258]

In the case of firearms examiners, the features they compared include class and individual characteristics. And it is the *comparison* of these features that forms the basis of their opinions.[259]

### *C. The Term "Foundational Validity" is not Recognized by the Legal or Scientific Communities*

Next, the PCAST Report created the term "foundational validity,"[260] which the PCAST Report claims is a scientific term.[261] Yet, the report again provides no scientific reference or citation for the definition except Federal Rule of Evidence 702, which is a legal source. Rule 702 only references "reliability."[262] Strangely, the words "foundational" and "validity" appear nowhere in the text of Rule 702. Neither of the U.S. Supreme Court's decisions in *Daubert* or *Kumho Tire* use the term "foundational validity."[263] It is an artificial term created and defined solely by PCAST.[264] It also differs from the term "scientific validity" espoused by the U.S. Supreme Court in

[258] Press Release, Dep't of Just., Just. Dep't Publishes Statement on 2016 President's Council of Advisors on Sci. & Tech. Rep. (Jan. 13, 2021), https://www.justice.gov/opa/pr/justice-department-publishes-statement-2016-presidents-council-advisors-science-and.

[259] *See AFTE Theory of Identification as it Relates to Toolmarks*, *supra* note 37.

[260] PCAST REPORT, *supra* note 10, at 4–5, 19, 43. PCAST has also defined "scientific validity" as when "a method has shown, based on empirical studies, to be reliable with levels of repeatability, reproducibility, and accuracy that are appropriate to the intended application." *Id.* at 48.

[261] *Id.* at 43.

[262] *Id.*; FED. R. EVID. 702.

[263] The Court used the terms "scientific validity" or "scientific knowledge," not "foundational validity," in these decisions. *See* Daubert v. Merrell Dow Pharms., Inc., 509 U.S. 579, 591, 594, 599–600 (1993); Kumho Tire Co. v. Carmichael, 526 U.S. 137, 141 (1999).

[264] Indeed, as Professor David Kaye of Penn State Law recognized, "'Foundational validity' is not a standard phrase in metrology and statistics." David H. Kaye, *Firearm-Mark Evidence: Looking Back and Looking Ahead*, 68 CASE W. RSRV. L. REV. 723, 737 n.77 (2018).

the *Daubert* decision, where the Court described the term as "trustworthiness."[265]

The International Organization for Standardization (ISO) is an independent, non-governmental international organization with a membership of 164 national standards bodies. It has published over 24,098 international standards on technology and manufacturing. One hundred sixty-five countries and 802 technical committees and subcommittees direct the ISO in the creation of these internationally accepted scientific standards.[266] In particular, ISO/IEC 17025 is the standard that pertains to the calibration and testing of laboratories, including those dealing in forensics.[267] Yet ISO 17025 neither recognizes nor uses PCAST's term "foundational validity," so the term's origin remains a mystery. Despite this lack of support, the PCAST Report recommends that expert testimony in "feature comparison methods," such as firearms identification, must first achieve "foundational validity" to be admitted in court.[268] The PCAST Report then attempts to extrapolate what it requires for a discipline to achieve "foundational validity."

It is important to dissect the key terms the PCAST Report uses to form the bases of the term "foundational validity." First is "repeatab[ility]," which PCAST defines as a "known probability, [that] an examiner obtains the same result, when analyzing samples from the same sources."[269] Next is "reproducib[ility]," which PCAST defines as "different examiners obtain[ing] the same result, when analyzing the same samples."[270] Given the existence of fifty validation studies in the field of firearms identification over the past century, one would be hard-pressed to claim the discipline has not demonstrated its results are either repeatable or reproducible. Nevertheless, this definition remains problematic because it is possible that two expert witnesses may use the same methodology, examine the same evidence, and reach totally different conclusions. Yet, taken to its logical destination, "reproducibility" could render expert testimony inadmissible where two

[265] 509 U.S. at 590 n.9 ("[O]ur reference here is to *evidentiary* reliability—that is, trustworthiness. . . . In a case involving scientific evidence, *evidentiary reliability* will be based upon *scientific validity*.").

[266] *See About Us*, INT'L ORG. FOR STANDARDIZATION, https://www.iso.org/about-us.html (last visited Jan. 14, 2022).

[267] INT'L ORG. FOR STANDARDIZATION, ISO/IEC 17025: GENERAL REQUIREMENTS FOR THE COMPETENCE OF TESTING AND CALIBRATION LABORATORIES (3d ed. 2017).

[268] PCAST REPORT, *supra* note 10, at 19.

[269] *Id.* at 47.

[270] *Id.*

experts offer different opinions regarding an examination of the same evidence. This is because a difference of opinion runs contrary to the concept of "reproducibility."

The U.S. Supreme Court recognized this inherent limitation on reproducibility in *Daubert* and counseled lower courts that "[t]he focus, of course, must be solely on principles and methodology, not on the conclusions that they generate."[271] The trial courts also do not view a lack of reproducibility or repeatability as a reason to question the reliability of a forensic discipline, even when expert witnesses may entertain different opinions as to identification or source attribution, such as DNA examination results in a criminal case:

> The Court recognizes that there are differences of expert opinion on source attribution. Admissibility under Rule 702 and *Daubert* does not require consensus, however, and the Court could conclude that this is an issue on which expert opinion on both sides is reliable enough for admission. A "battle of experts" is for the jury to resolve.[272]

The next definition concerns "accura[cy]," which PCAST claims exists when the "examiner obtains correct results . . . for samples from the same source (true positives)" and "samples from different sources (true negatives)."[273] Note that "accuracy" contains two components. This raises an unanswered question: could a discipline have high false negatives, but few—if any—false positives, and still be deemed accurate?

Beyond the key terms cited above, the PCAST Report's requirements for "foundational validity" go on to mandate no less than six separate requirements, all of which must be utilized within a forensic discipline for it

[271] 509 U.S. 579, 595 (1993).

[272] United States v. McCluskey, 954 F. Supp. 2d 1224, 1269 (D.N.M. 2013). The *McCluskey* case entailed the admissibility of a DNA expert testifying for the prosecution. While the court found the expert witness could testify, she was not permitted to testify about "low copy number" from "touch" DNA. *Id.* at 1230, 1276, 1292. Firearms and DNA are not the only disciplines or sciences where experts can reach contrary opinions when examining the same data. The U.S. Supreme Court observed in an unrelated context, "Psychiatry is not . . . an exact science, and psychiatrists disagree widely and frequently on what constitutes mental illness, on the appropriate diagnosis to be attached to given behavior and symptoms, on cure and treatment, and on likelihood of future dangerousness." Ake v. Oklahoma, 470 U.S. 68, 81 (1985).

[273] PCAST REPORT, *supra* note 10, at 47.

to achieve foundational validity.[274] Yet these rigid, dogmatic criteria for admissibility of expert testimony stand inapposite to the U.S. Supreme Court's decision in *Kumho Tire*, where the Court emphasized that the application of Rule 702 is "a flexible one."[275] Regardless of the merits of the PCAST's criteria for "foundational validity," nothing mandates that any or all of them are a prerequisite to admissibility of expert testimony.

Chief among these requirements is the use of *only* "black box" studies[276] to demonstrate the reliability of a science or discipline. The PCAST Report defines a "black box" study as "an empirical study that assesses a subjective method by having examiners analyze samples and render opinions about the

[274]Ted Robert Hunt, *Scientific Validity and Error Rates: A Short Response to the PCAST Report*, 86 FORDHAM L. REV. ONLINE 24, 27 (2018), https://fordhamlawreview.org/wp-content/uploads/2018/03/Hunt_DOJ-24.pdf. Hunt lists the nine requirements for "appropriately designed studies" as outlined by the PCAST Report, which the author has distilled into six criteria as follows:

> "(1) The studies must involve a sufficiently large number of examiners and must be based on sufficiently *large* collections of *known* and *representative* samples from *relevant* populations to reflect the range of features or combinations of features that will occur in the application. . . .
>
> (2) The empirical studies should be conducted so that neither the examiner nor those with whom the examiner interacts have any information about the correct answer.
>
> (3) The study design and analysis framework should be specified in advance. In validation studies, it is inappropriate to modify the protocol afterwards based on the results.
>
> (4) The empirical studies should be conducted or overseen by individuals or organizations that have no stake in the outcome of the studies.
>
> (5) Data, software and results from validation studies should be available to allow other scientists to review the conclusions.
>
> (6) To ensure that conclusions are reproducible and robust, there should be multiple studies by separate groups reaching similar conclusions." *Id.*

[275]526 U.S. 137, 150 (1999). *But see id.* at 159 (Scalia, J., concurring) ("[I]t is discretion to choose among *reasonable* means of excluding expertise that is *fausse* and science that is junky. Though, as the Court makes clear today, the *Daubert* factors are not holy writ . . . .").

[276]PCAST REPORT, *supra* note 10, at 46. PCAST uses the term "black box" ninety-one times in the report and the term "black box studies" fifty-five times. By comparison, the terms "black box study" and "white box" appear nowhere in either NRC's BALLISTIC IMAGING, *supra* note 11, from 2008, or the NRC REPORT, *supra* note 12, from 2009.

origin or similarity of samples."[277] These studies must deliver "a reproducible and consistent procedure for . . . identifying features within evidence samples" and derive "empirical measurements, from multiple independent studies."[278] This means two or more such studies are required.[279] Nevertheless, PCAST cannot cite any legal or scientific basis for this requirement, other than its own interpretation of "foundational validity," which mandates a particular methodology be "repeatable, reproducible, and accurate."[280] Again, neither Rule 702 nor the U.S. Supreme Court's decisions in *Daubert* or *Kumho Tire* require such criteria or factors to assess the reliability of expert witness testimony or the discipline about which the experts testify. The National Research Council does not recognize the distinction between using "black box" studies or "white box"[281] studies to establish an error rate.[282] The NRC's 2009 *Strengthening Forensic Science Report* never uses the term "black box study" or "black box studies."[283] Nor does the NRC distinguish between closed or open-set studies in assessing the reliability of firearms identification as a discipline.

### *D. Reliance on the NRC's 2008 Ballistic Imaging and 2009 Strengthening Forensic Science Reports*

To support its attack of the firearms identification discipline, the PCAST Report cited both the 2008 *Ballistic Imaging Report* and the 2009 *Strengthening Forensic Science Report* from the National Research Council (NRC).[284] However, by 2016, the analysis contained in these two reports was either inapplicable or already stale for four reasons. For starters, the 2008 *Ballistic Imaging Report* is a poor reference for PCAST or any other critic of

[277] PCAST REPORT, *supra* note 10, at 48. "In black-box studies, many examiners are presented with many independent comparison problems—typically, involving 'questioned' samples and one or more 'known' samples—and asked to declare whether the questioned samples came from the same source as one of the known samples. The researchers then determine how often examiners reach erroneous conclusions." *Id.* at 49 (footnote omitted).

[278] *Id.* at 48.

[279] *See id.*

[280] *Id.* at 4, 47 (emphasis omitted).

[281] A "white box" study examines the underlying process employed by a forensic examiner and factors that affect the examiner's decisions, though it does not analyze the examiner. *See id.* at 9.

[282] *See* NRC REPORT, *supra* note 12.

[283] *Id.*

[284] PCAST REPORT, *supra* note 10, at 104–05.

the firearms identification discipline to rely upon because, as the NRC expressly stated:

> First, and most significantly, *this study is neither a verdict on the uniqueness of firearms-related toolmarks generally nor an assessment of the validity of firearms identification as a discipline. . . . [T]he proposal for this study explicitly precluded the committee from assessing the admissibility of forensic firearms evidence in court*, either generally or in specific regard to testimony on ballistic imaging comparisons.[285]

Second, PCAST's reliance on the 2009 NRC *Strengthening Forensic Science Report* is similarly misplaced. As the courts have noted, the "purpose of the NAS Report is to highlight deficiencies in a forensic field and to propose improvements to existing protocols, not to recommend against admission of evidence."[286] Indeed, the project's co-chair, Judge Harry Edwards, made it clear that nothing in the report was intended to answer the "question whether forensic evidence in a particular case is admissible under applicable law."[287] Judge Edwards also told Congress that "the report offers no proposals for law reform. That was beyond our charge."[288]

---

[285] BALLISTIC IMAGING, *supra* note 11, at 18, 20. *See also* Affidavit of Dr. John E. Rolph, United States v. Edwards, No. F-516-01 (D.C. Super. Ct. 2008), https://afte.org/uploads/documents/swggun-rolph-affidavit.pdf. Rolph, a statistics professor at the University of Southern California, served as the chair for the committee that wrote and published the NRC's 2008 *Ballistic Imaging Report*. *Id.* at 1. It should be noted the *Ballistic Imaging Report*'s duty was to opine on the feasibility of a ballistics database, not whether forensic examiners could identify bullets or cartridge cases as coming from a particular firearm. BALLISTIC IMAGING, *supra* note 11, at 2.

[286] State v. McGuire, 16 A.3d 411, 436 (N.J. Super. Ct. App. Div. 2011).

[287] United States v. Rose, 672 F. Supp. 2d 723, 725 (D. Md. 2009).

[288] *Strengthening Forensic Science in the United States: A Path Forward: Hearing Before the S. Comm. on the Judiciary*, 111th Cong. 10 (2009) [hereinafter *Forensic Science in the United States*] (statement of Hon. Harry T. Edwards). Judge Edwards also remarked, "[t]he findings and recommendations of the committee do not mean to offer any judgments on any cases in the judicial system. . . . [E]ach case in the criminal justice system must be decided on the record before the court pursuant to the applicable law, controlling precedent, and governing rules of evidence. The question whether forensic evidence in a particular case is admissible under applicable law is not coterminous with the question whether there are studies confirming the scientific validity and reliability of a forensic science discipline." *Id.* This last sentence might well have been a warning to the authors of the PCAST Report.

Third, many of the recommendations in the 2009 NRC *Strengthening Forensic Science Report*, such as accreditation, quality controls, and proficiency testing,[289] were adopted by an overwhelming majority of forensic laboratories.[290]

Fourth, at least six new firearms studies were published in peer-reviewed journals since the two NRC reports were published, answering the NRC's call for additional studies in the firearms field.[291] Together, these six studies consistently demonstrate remarkably low rates of false identification, typically one percent or less.[292] They also demonstrate that firearms identification is repeatable, reproducible, accurate, and reliable.

## *E. PCAST Gets the Numbers on Firearms Studies Wrong*

PCAST addressed those six firearms studies and misconstrued the numbers in two of those studies. PCAST cites the 2014 black box *Baldwin (Ames) Study* as having a false-positive rate of 1.5% and a 95% confidence-bound interval of 2.2%.[293] Yet the *Baldwin (Ames) Study* itself reports a false-positive rate of just 1.01% and a 95% confidence-bound interval[294] ranging

[289] NRC REPORT, *supra* note 12, at 215.

[290] BUREAU OF JUST. STATS., OFF. OF JUST. PROGRAMS, U.S. DEPT. OF JUST., NCJ 250152, PUBLICLY FUNDED FORENSIC CRIME LABORATORIES: QUALITY ASSURANCE PRACTICES, 2014, at 1 (2016) ("In 2014, 88% of the nation's crime labs were accredited by a professional organization, up from 70% in 2002," and, "[a]s in previous years, nearly all (98%) crime labs conducted proficiency testing in 2014.").

[291] *See Hamby Study*, *supra* note 24; *Fadul Study*, *supra* note 192; *Stroman Study*, *supra* note 206; *Miami-Dade Study*, *supra* note 184; *Baldwin Study*, *supra* note 211; *Smith Study*, *supra* note 221.

[292] A seventh firearms study was published in 2017. *See* Mark A. Keisler et al., *Isolated Pairs Research Study*, 50 AFTE J. 56 (2017). Keisler had 126 different examiners compare cartridge cases from nine different SW .40 pistols. *Id.* at 56–57. The results found that examiners correctly identified the firearms which fired the cartridge cases in 1,508 of 1,512 instances, with no false negatives or false positives, creating a zero percent error rate. *Id.* at 57. Inconclusive results were not counted as errors. *Id.* at 56. This study was peer reviewed. *Id.*

[293] PCAST REPORT, *supra* note 10, at 111 tbl.2 (referencing the "Ames Laboratory Study"). The *Baldwin (Ames) Study* was conducted by David P. Baldwin at the Ames Laboratory on behalf of the Defense Forensic Science Center. It studied 218 firearms examiners who analyzed cartridge cases. The study found they had a false identification rate of 1.01% and a false exclusion rate of 0.36%. *Baldwin Study*, *supra* note 211, at 17 tbl.III.

[294] A 95% "frequency" or "confidence bound" interval is the rate at which one would expect to see false positive results 95% of the time.

from 0.36% to 2.26%.[295] Both numbers were substantially less than those reported by PCAST.

In the 2013 Fadul or Miami-Dade studies, PCAST claimed a 2.1% false-positive rate with a 95% confidence-bound interval of 4.7%, which equates to one false-positive result for every twenty-one firearms examinations.[296] Yet the *Miami-Dade Study* reported an error rate of just 0.7% and a 95% confidence-bound interval, with an upper-bound error rate of only 1.2%.[297] PCAST appears to have artificially inflated the cited error rates by ignoring inconclusive results for unmatched sets when calculating its false-positive error rate. Consequently, the PCAST error-rate calculations for the *Miami-Dade Study* are significantly overstated and misleading. Unfortunately, some courts have cited PCAST's inaccurate numbers as a basis to exclude or curtail testimony of firearms-identification experts due to an unacceptably high error rate in the firearms discipline.[298]

### *F. The PCAST Report is Not Peer Reviewed*

One of the five *Daubert* factors highlighted the publication of studies in peer-reviewed journals as a means to evaluate a science, discipline, or technical field.[299] The PCAST Report also places great emphasis on forensic studies being "peer reviewed" in scientific publications. PCAST identifies itself as a scientific entity and issued a report based on science, yet the PCAST Report itself is not peer reviewed,[300] and nothing in the report indicates that it was ever subjected to peer review by any scientific journal. It should come as no surprise that a chorus of more than a dozen organizations issued statements in response to the PCAST Report.[301] Many of these

[295] *Baldwin Study*, *supra* note 211, at 17 tbl.III.

[296] PCAST REPORT, *supra* note 10, at 109, 111 tbl.2 (referencing *Miami-Dade Study*).

[297] *Miami-Dade Study*, *supra* note 184, at 33.

[298] *See* United States v. Adams, 444 F. Supp. 3d 1248, 1264–65 (D. Or. 2020); United States v. Shipp, 422 F. Supp. 3d 762, 778 (E.D.N.Y. 2019).

[299] 509 U.S. 579, 593–94 (1993).

[300] United States v. Tibbs, No. 2016-CF1-19431, 2019 D.C. Super. LEXIS 9, at *53 (D.C. Super. Ct. 2019). Ironically, the term "peer review" appears some fifteen times in the PCAST Report and is considered a factor in determining whether a study or report is deemed reliable. PCAST REPORT, *supra* note 10, at 11, 32, 55, 66, 95, 111, 125, 129, 147.

[301] The following organizations released statements in response to the PCAST Report: American Academy of Forensic Sciences; American Congress of Forensic Science Laboratories; American Society of Crime Lab Directors: Association of Firearm and Tool Mark Examiners; Bureau of Alcohol, Tobacco, Firearms and Explosives; Department of Justice; Federal Bureau of

organizations sharply criticized PCAST's critique of firearms identification, including the American Society of Crime Laboratory Directors;[302] the AFTE;[303] the Organization of Scientific Area Committee (OSAC) Firearms and Toolmarks Subcommittee;[304] the Bureau of Alcohol, Tobacco, Firearms and Explosives (ATF);[305] and the FBI.[306] On behalf of the DOJ, Attorney General Lorretta Lynch—the legal advisor to President Barack Obama—issued the following statement, rejecting the findings and recommendations of the PCAST Report:

Investigation; Forensic Institute; Innocence Project; International Association for Identification; Midwestern Association of Forensic Scientists; National Association of Criminal Defense Lawyers; and the National District Attorneys Association. *Published Statements in Response to the PCAST Report on Forensic Science in Criminal Courts*, OBAMA WHITE HOUSE ARCHIVES, https://obamawhitehouse.archives.gov/sites/default/files/microsites/ostp/PCAST/pcast_forensics_2016_public_comments.pdf (last visited Mar. 9, 2020) (publishing a complete list of organizations releasing statements along with links to each statement). PCAST responded to many of these critical statements in a nine-page addendum from January 6, 2017. PRESIDENT'S COUNCIL OF ADVISORS ON SCI. & TECH., AN ADDENDUM TO THE PCAST REPORT ON FORENSIC SCIENCE IN CRIMINAL COURTS (2017), https://obamawhitehouse.archives.gov/sites/default/files/microsites/ostp/PCAST/pcast_forensics_addendum_finalv2.pdf.

[302] Statement, Am. Soc'y of Crime Lab'y Dirs., Inc., Statement on September 20, 2016 PCAST Rep. on Forensic Sci. (Sept. 30, 2016), https://pceinc.org/wp-content/uploads/2016/10/20160930-Statement-on-PCAST-Report-ASCLD.pdf.

[303] Statement, Ass'n of Firearms & Tool Mark Exam'rs, Response to PCAST Report on Forensic Sci. (Oct. 31, 2016), https://afte.org/uploads/documents/AFTE-PCAST-Response.pdf.

[304] Statement, Org. of Sci. Area Comms., Response to the President's Council of Advisors on Sci. & Tech. (PCAST) Call for Additional References Regarding its Report "Forensic Sci. in Crim. Cts.: Ensuring Sci. Validity of Feature-Comparison Methods" (Dec. 14, 2016), https://www.nist.gov/system/files/documents/2016/12/16/osac_firearms_and_toolmarks_subcommittees_response_to_the_presidents_council_of_advisors_on_science_and_technologys_pcast_request_for_additional_references_-_submitted_december_14_2016.pdf. The Organization of Scientific Area Committees (OSAC) is sponsored by the National Institute of Standards and Technology (NIST) under the control of the Department of Commerce. The Firearm and Toolmark Subcommittee has sixteen forensic examiners, with a combined 307 years of forensic science experience.

[305] Statement, Bureau of Alcohol, Tobacco, Firearms and Explosives, ATF Response to the President's Council of Advisors on Sci. & Tech. Rep. (Sept. 11, 2016), https://theiai.org/docs/9.20160921_ATF_PCAST_Response.pdf.

[306] Statement, Fed. Bureau of Investigation, Comments on: President's Council of Advisors on Sci. & Tech. Rep. to the President, Forensic Sci. in Fed. Crim. Cts.: Ensuring Sci. Validity of Pattern Comparison Methods (Sept. 20, 2016), www.fbi.gov/file-repository/fbi-pcast-response.pdf.

> We remain confident that, when used properly, forensic science evidence helps juries identify the guilty and clear the innocent, and the department believes that the current legal standards regarding the admissibility of forensic evidence are based on sound science and sound legal reasoning. While we appreciate their contribution to the field of scientific inquiry, the department will not be adopting the recommendations related to the admissibility of forensic science evidence.[307]

### *G. The PCAST Report Erroneously Claims Casework and Experience do not Constitute a Basis for the Admission of Expert Testimony*

Finally, the PCAST Report declares that "[c]asework [alone] is not scientifically valid research, and experience alone cannot establish scientific validity" as a condition precedent to the admission of an expert's testimony.[308] This is flatly untrue. The Comments to the 2000 Amendment to Rule 702 made the following clear:

> Nothing in this amendment is intended to suggest that experience alone—or experience in conjunction with other knowledge, skill, training or education—may not provide a sufficient foundation for expert testimony. To the contrary, the text of Rule 702 expressly contemplates that an expert may be qualified on the basis of experience. In certain fields, experience is the predominant, if not sole, basis for a great deal of reliable expert testimony.[309]

---

[307] Gary Fields, *White House Advisory Council Report Is Critical of Forensics Used in Criminal Trials*, THE WALL STREET JOURNAL (Sept. 20, 2016, 4:25 PM), https://www.wsj.com/articles/white-house-advisory-council-releases-report-critical-of-forensics-used-in-criminal-trials-1474394743.

[308] PCAST REPORT, *supra* note 10, at 32–33.

[309] FED. R. EVID. 702 advisory committee's note to 2000 amendment. Rule 702 provides: "Testimony by Expert Witnesses. A witness who is qualified as an expert by knowledge, skill, *experience*, training, or education may testify in the form of an opinion or otherwise if:

(a) the expert's scientific, *technical, or other specialized knowledge* will help the trier of fact to understand the evidence or to determine a fact in issue;

(b) the testimony is based on sufficient facts or data;

(c) the testimony is the product of reliable principles and methods; and

Whether a discipline is a "science" or the witness is a "scientist" is not the litmus test to determine the reliability or admissibility of expert witness testimony. Experience remains a powerful reason to admit expert witness testimony. Experience also matters in the eyes of the U.S. Supreme Court, as it found in *Kumho Tire v. Carmichael*:

> *Daubert* pointed out that Federal Rules 702 and 703 grant expert witnesses testimonial latitude unavailable to other witnesses on the "assumption that the expert's opinion will have a reliable basis in the knowledge and experience of his discipline." The Rules grant that latitude to all experts, not just to "scientific" ones.[310]

The PCAST report is neither a scientifically nor legally reliable report on which to base any decision governing the admissibility of firearms expert testimony. This is especially true when one considers the PCAST Report's questionable use of terms such as "metrology" and "foundational validity," the absence of peer review, the lack of anyone on the PCAST staff with any firearms examination or forensic science experience, and the inaccurate reporting of the actual error rate in firearm identification studies.[311]

## IX. POST-PCAST EVIDENTIARY RULINGS ON FIREARMS ANALYSIS EXPERT TESTIMONY, 2016–PRESENT

Since the release of the PCAST report in 2016, several state and federal courts have reviewed the admissibility of firearms identification expert witness testimony. Most of these jurisdictions continue to uphold the

(d) the expert has reliably applied the principles and methods to the facts of the case." FED. R. EVID. 702 (emphasis added).

[310] 526 U.S. 137, 148 (1999) (citation omitted).

[311] One question which arises from the PCAST Report is whether a firearms identification expert witness can be cross examined on the contents of the PCAST Report. The PCAST Report is hearsay under Federal Rule of Evidence 801 and is not admissible under Federal Rule of Evidence 802, so it would not be admissible in most cases. However, Federal Rule of Evidence 803(18) permits an exception for "learned treatises, periodicals, or pamphlets." FED. R. EVID. 803(18). If neither an examiner for the prosecution nor the defense testifies the PCAST Report is considered as reliable in the field, then the Report cannot be read into evidence unless the court takes judicial notice that the report is considered reliable. Given the myriad of problems with the PCAST Report, its rejection by the Department of Justice, and vocal criticism of the Report within the forensic science community, courts would be wise to avoid any judicial notice finding that the PCAST Report is a reliable authority in the area of firearms identification expert testimony, or any other forensic discipline.

admissibility of firearms identification experts to testify and to do so without restrictions or limitations on their testimony.

In *United States v. Johnson*, the U.S. Court of Appeals for the Ninth Circuit affirmed a conviction based upon firearms identification expert testimony that "matched" a bullet recovered from the crime scene to a pistol found in the possession of the defendant.[312] The court acknowledged the questions raised by the 2009 National Academy of Sciences Report regarding the firearms discipline, including criticism of the AFTE Theory of Identification.[313] While the court noted a number of cases where experts were precluded from testifying that any match was an "absolute certainty," it noted that the defendant could find only one case[314] where the court did not permit a firearms identification expert to testify as to a match.[315] The court also acknowledged that the expert in *Johnson* was not "absolutely certain" in his testimony, that he was subject to cross examination by the defense, and that the defense was free to call its own expert at trial.[316] The court found these were "adequate safeguards" used by the district court when admitting firearms identification expert testimony.[317] The court recognized that the district court cited numerous cases where the AFTE Theory of Identification satisfied *Daubert*, while the defendant did not cite a single case where AFTE ballistics testimony had been excluded altogether.[318]

Another federal appellate court reached a similar result in *United States v. Gil*, where the U.S. Court of Appeals for the Second Circuit affirmed a district court judge's decision to admit the unrestricted testimony of a firearms identification expert witness.[319] The court acknowledged an error rate for the ballistics identification technique "in the range of 1%," which the

[312] 875 F.3d 1265, 1279–80 (9th Cir. 2017).

[313] *Id.* at 1280.

[314] United States v. Glynn, 578 F. Supp. 2d 567, 574–75 (S.D.N.Y. 2008) (allowing the expert to testify only that it was "more likely than not" that bullets matched).

[315] *Johnson*, 875 F.3d at 1280.

[316] *Id.* at 1280–81.

[317] *Id.* at 1281.

[318] *Id.* The Ninth Circuit does not mention the PCAST report in its 2017 opinion. As the defendant's trial occurred in early 2016 and the PCAST report was released later in the year, it is likely the PCAST report was not examined by the district court.

[319] 680 Fed. App'x. 11 (2d Cir. 2017). While the 2008 and 2009 NAS Reports on Firearms are referenced by the Second Circuit in its decision, the PCAST Report is not mentioned as the conviction occurred several months before the PCAST Report was released in late 2016.

court characterized as a "de minimis" potential rate of error.[320] The court threw its unequivocal support behind the firearms identification discipline, holding the appellant's "challenges to the admission of ballistics expert opinion are meritless."[321] Additionally, the court held that "arguments about the subjectivity inherent in otherwise reliable methodologies go 'to weight of the evidence, not to its admissibility,' and were 'matters for cross-examination and argument to the jury.'"[322]

In *United States v. Brown*, the U.S. Court of Appeals for the Seventh Circuit also reviewed and affirmed a conviction on the basis of firearms identification expert testimony in which the firearms experts explicitly testified that the cartridge cases found at multiple murder scenes were a match.[323] The accused objected to the admission of such testimony, claiming that the PCAST Report and the lack of objective, quantifiable standards to determine whether there exists a match between ammunitions components, rendered firearms expert testimony unreliable.[324] The court rejected these arguments, finding that the district court did not abuse its discretion in admitting the firearms identification expert testimony.[325] The court found "[a]lmost all the defendants' contentions were issues that could be raised on cross-examination. These arguments go to the weight of the evidence, not its admissibility. Expert testimony is still testimony, not irrefutable fact, and its ultimate persuasive power is for the jury to decide."[326]

In *Garrett v. Commonwealth*, the Kentucky Supreme Court faced another post-PCAST challenge to firearms identification expert testimony.[327] Here, the firearms expert identified the pistol obtained from the defendant as having fired the bullet recovered during a murder investigation.[328] The defendant relied on the 2009 NRC Report to claim the AFTE Theory of Identification utilized by the firearms expert was not reliable, thus rendering the firearm

[320] *See id.* at 13.

[321] *Id.* at 14.

[322] *Id.* (quoting United States v. Romano, 794 F.3d 317, 333 (2d Cir. 2015)).

[323] 973 F.3d 667, 702–04 (7th Cir. 2020), *cert. denied*, 141 S. Ct. 1253 (2021); *see also* United States v. Godinez, 7 F.4th 628, 633–36 (7th Cir. 2021) (upholding another district court's admission of firearms identification expert testimony).

[324] *Brown*, 973 F.3d at 703.

[325] *Id.* at 704.

[326] *Id.*

[327] 534 S.W.3d 217, 222 (Ky. 2017).

[328] *See id.*

expert's testimony inadmissible.[329] The Kentucky Supreme Court rejected this argument, finding that the AFTE Theory of Identification satisfied the *Daubert* factors and holding that the trial court did not abuse its discretion in admitting such testimony.[330] In addition, the court held the following:

> The proper avenue for Garrett to address his concerns about the methodology and reliability of Collier's testimony was through cross-examination, as well as through the testimony of his own expert. In this way, the jury was presented with both parties' positions, and with any limitations to the testimony, and charged with weighing all the evidence presented.[331]

Since the PCAST Report's release in 2016, the list of state courts reaffirming the admission of firearms identification expert witness testimony at trial continues to grow: appellate courts in California,[332] Connecticut,[333] Delaware,[334] Mississippi,[335] Washington,[336] North Carolina,[337] Louisiana,[338] Maryland,[339] Missouri,[340] Nebraska,[341] New Jersey,[342] New Mexico,[343] and

[329] *Id.*

[330] *Id.* at 222–23.

[331] *Id.* at 223.

[332] People v. Therman, No. C091147, 2021 WL 4859299, at *1 (Cal. Ct. App. Oct. 19, 2021).

[333] State v. Raynor, 189 A.3d 652, 656 (Conn. App. Ct. 2018); State v. Terrell, No. CR170179563, 2019 WL 2093108, at *1 (Conn. Super. Ct. Mar. 21, 2019).

[334] State v. Mobley, ID No. 2002007105, 2021 WL 5411089, at *1–2 (Del. Super. Ct. Nov. 19, 2021).

[335] Willie v. State, 274 So. 3d 934, 935 (Miss. Ct. App. 2018).

[336] State v. DeJesus, 436 P.3d 834, 837–38 (Wash. Ct. App. 2019); State v. Hatfield, No. 77512-0-I, 2019 WL 6492483, at *8–9 (Wash. Ct. App. Dec. 2, 2019).

[337] State v. Williams, 814 S.E.2d 925 (N.C. Ct. App. 2018); State v. Griffin, 834 S.E.2d 435, 436 (N.C. Ct. App. 2019); State v. Miller, 852 S.E.2d 704, 706 (N.C. Ct. App. 2020).

[338] State v. Lee, 217 So. 3d 1266, 1278 (La. Ct. App. 2017); *see also* State v. Magee, 243 So. 3d 151 (La. Ct. App. 2018).

[339] Patterson v. State, 146 A.3d 496, 497 (Md. Ct. Spec. App. 2016); Abruquah v. State, No. 2176, 2020 WL 261722, at *1 (Md. Ct. Spec. App. Jan. 17, 2020).

[340] State v. Boss, 577 S.W.3d 509, 512 (Mo. Ct. App. 2019); State v. Mills, 623 S.W.3d 717 (Mo. Ct. App. 2021). These cases superseded the trial court case of *Missouri v. Goodwin-Bey*, No. 1531-CR00555-01 (Mo. Cir. Ct. Dec. 16, 2016).

[341] State v. Wheeler, 956 N.W.2d 708, 719 (Neb. 2021).

[342] State v. Oliver, No. A-5140-16T1, 2020 WL 773578, at [*1] (N.J. Super. Ct. App. Div. Feb. 18, 2020).

[343] State v. Nowicki, No. S-1-SC-37388, 2020 WL 1910847, at *1 (N.M. Apr. 20, 2020).

Ohio[344] have all reviewed firearms identification expert testimony in the shadow of the PCAST Report and found such expert testimony admissible, with few, if any, restrictions. Only one state court—New York—has precluded a firearms expert from expressing any opinion as to the identification of a cartridge case or bullet.[345]

Federal courts have also weighed in on the admissibility of firearms identification expert testimony. Since the PCAST report was released, U.S. district courts in Arizona,[346] California,[347] the District of Columbia,[348] Nevada,[349] New York,[350] Oklahoma,[351] and Virginia[352] also found firearms identification expert testimony admissible without restrictions. Many courts flatly rejected the findings of the PCAST Report:

> [T]he PCAST report acknowledged its own dubious value to courts, stating, "Judges' decisions about the admissibility of scientific evidence rest solely on legal standards; they are exclusively the province of the courts and PCAST does not opine on them." . . . [T]he PCAST report here does not indicate that the toolmark testing is without merit. Instead, it urges experts to use certain approaches and methodology. . . . [T]he reports on which DeJesus [the defendant] relies do not affect the general scientific acceptance of ballistic identification. Instead, the problems they espouse bear on the question of reliability of the individual test and tester at issue. These questions are then

[344] State v. Smith, No. 109402, 2021 WL 507706, at *3 (Ohio Ct. App. 2021).

[345] People v. Ross, 129 N.Y.S. 3d 629, 641–42 (N.Y. Sup. Ct. 2020).

[346] Merritt v. Arizona, No. CV-17-04540-PHX-DGC, 2021 WL 1541635, at *3 (D. Ariz. Apr. 20, 2021).

[347] United States v. Chavez, No. 15-CR-00285-LHK-1, 2021 WL 5882466, at *17–18 (N.D. Cal. Dec. 13, 2021).

[348] United States v. Harris, 502 F. Supp. 3d 28, 33 (D.D.C. 2020).

[349] United States v. Romero-Lobato, 379 F. Supp. 3d 1111, 1114 (D. Nev. 2019).

[350] United States v. Johnson, (S5) 16 Cr. 281 (PGG), 2019 WL 1130258, at *1–2 (S.D.N.Y. Mar. 11, 2019).

[351] United States v. Hunt, 464 F. Supp. 3d 1252, 1262 (W.D. Okla. 2020).

[352] United States v. Simmons, No. 2:16cr130, 2018 U.S. Dist. LEXIS 18606, at *4 (E.D. Va. 2018).

> considered by the trier of fact in assessing the weight to be given the evidence.[353]

Yet the courts have not achieved unanimity when it comes to the admissibility of firearms identification expert testimony. A small minority of state and federal cases have severely restricted firearms examiner testimony.[354] Although none of these decisions denied the admission of a firearms identification expert witness's testimony, these decisions severely limited—or even re-wrote—these experts' testimony.[355] The courts precluded the experts from identifying a particular firearm as the source of a questioned cartridge case or bullet, effectively nullifying the experts' testimony.[356] All courts relied heavily on the flawed PCAST Report to justify

[353] State v. DeJesus, 436 P.3d 834, 841–42 (Wash. Ct. App. 2019) (citations omitted) (emphasis in original).

[354] *See* Williams v. United States, 210 A.3d 734, 744 (D.C. 2019) (holding that "it is plainly error to allow a firearms and toolmark examiner to unqualifiedly opine, based on pattern matching, that a specific bullet was fired by a specific gun"); *see also* United States v. Tibbs, No. 2016 CF1 19431, 2019 D.C. Super. LEXIS 9, at *85 n.18 (D.C. Super. Ct. 2019) (ordering the expert to testify only that "the recovered firearm cannot be excluded as the source of the recovered casing" found at the crime scene in lieu of his opinion of "identification"); *see also* United States v. Davis, No. 4:18-cr-00011, 2019 WL 4306971, at *7 (W.D. Va. 2019) (instructing the expert to only testify as to if the marks on the cartridge cases from the crime scene were "consistent with" those from the defendant's gun, instead of a "match" or identification); *see also* United States v. Shipp, 422 F. Supp. 3d 762, 783 (E.D.N.Y. 2019) (ordering the expert to testify that the marks on the bullet fragment and shell casing were "consistent with" those made by the defendant's gun and that the defendant's gun "cannot be excluded as the source of the bullet fragment or cartridge case" in lieu of identifying the defendant's gun as the source of the bullet fragment or cartridge case); *see also* Transcript of Record at 119–20, 126, 130, United States v. Medley, 312 F. Supp. 3d 493 (D. Md. 2018) [hereinafter Medley Transcript] (permitting the expert to testify that the marks on the cartridge case found at the crime scene were "consistent with" those made by a cartridge case fired by the defendant's gun, in lieu of identifying the defendant's weapon as the source of the fired cartridge case); *see also* United States v. Adams, 444 F. Supp. 3d 1248, 1267 (D. Or. 2020) (permitting the expert to testify only about class characteristics of the bullet recovered from the crime scene and those of the bullet fired by the defendant's gun, in lieu of testifying about individual characteristics and that the bullets were a "match"); *see also* People v. Ross, 129 N.Y.S.3d 629, 630, 642 (N.Y. Sup. Ct. 2020) (permitting the firearms expert to testify only about class characteristics and nothing more, in lieu of a "match" of the individual characteristics of the shell casings found at the crime scene to ones made by a gun found in the defendant's car).

[355] *See DeJesus*, 436 P.3d at 841–42. While the court ultimately affirmed the admission of firearms identification expert testimony, it catalogued those opinions which placed restrictions on such expert testimony.

[356] *See, e.g.*, *Williams*, 210 A.3d at 744.

their decisions.[357] In doing so, they failed to properly apply Rule 702, as well as the decisions in *Daubert v. Merrell Dow Pharmaceuticals*[358] and *Kumho Tire v. Carmichael*.[359] Their decisions to rewrite the testimony of firearms identification expert witnesses—or to deny such testimony altogether—constitutes an abuse of discretion by the judges who made these flawed decisions.[360]

## X. THE FAILURE OF THE POST-PCAST CASES THAT RESTRICT OR DENY THE ADMISSIBILITY OF FIREARMS IDENTIFICATION EXPERT TESTIMONY

> *"[T]he danger is, and I think there's already some of this happening, but the danger is that there is going to be a broader exclusion of legitimate evidence, because I think the judges will think that their job is not just to decide, by preponderance of the evidence, whether it's reliable, but you're going to see judges who say I have to decide if your science is right. There's a lot of that going on."*
>
> Circuit Judge Katherine M. O'Malley, U.S. Court of Appeals for the Federal Circuit[361]

We begin with one of the first post-PCAST Report cases to restrict or re-write the opinion of a firearms identification expert witness: the 2018 case of *United States v. Medley*.[362] The judge presiding over the trial and the admissibility hearing[363] was no stranger to the admissibility of firearms identification expert witness testimony, having served as a magistrate judge

[357] *See, e.g., id.* at 741.

[358] 509 U.S. 579 (1993).

[359] 526 U.S. 137 (1999).

[360] Under the U.S. Supreme Court's decision in *Gen. Elec. v. Joiner*, a trial court's decision whether to admit evidence or testimony of an expert witness is subject to review by the appellate court. 522 U.S. 136, 143 (1997). The appellate court may overturn the trial court's decision if it finds the trial court abused its discretion. *Id.*

[361] *Conference on Proposed Amendments: Experts, the Rule of Completeness, and Sequestration of Witnesses*, 87 FORDHAM L. REV. 1361, 1376 (2019) (speaking at the Philip D. Reed Lecture Series, Advisory Committee on Evidence Rules).

[362] Medley Transcript, *supra* note 354. The first post-PCAST Report case to restrict or modify a firearms expert witness testimony was *Missouri v. Goodwin-Bey*, No. 1531-CR00555-01 (Mo. Cir. Ct. Dec. 16, 2016). *See supra* note 340.

[363] Medley Transcript, *supra* note 354, at 1.

in the case of *United States v. Willock*, which dealt with the same issue eight years prior.[364] In *Willock*, the judge permitted Professor Adina Schwartz to testify about firearms examination expert witness testimony and relied extensively on her testimony in reaching his conclusions.[365] He then recommended changing the firearm examiner's testimony in a number of particulars. Essentially the examiner could no longer declare that two cartridge cases found at different murder crime scenes were a match.[366] Instead, he could only state it was "more likely than not" that the two cartridge cases were a match.[367]

Returning to the court's 2018 *Medley* decision, the court's analysis suggests the possibility of confirmation bias. For example, the judge cited to his previous 2009 decision in the *Willock* case no fewer than twenty-two times.[368] The court ruled a firearms identification expert—who had testified he could match a bullet to the firearm in the case—would only be permitted to say the marks on the bullet were "consistent with" those made by the firearm.[369] The judge forbade the examiner from testifying that the cartridge cases found at the crime scene were fired by the gun associated with the defendant and prohibited the examiner from using the terms "identify" or "identification."[370] Instead the court substituted the term "consistent with" for the examiner's term of "identify."[371]

---

[364] 696 F. Supp. 2d 536 (D. Md. 2010).

[365] *Id.* at 567–68. This was the *same* Adina Schwartz who the U.S. District Court in New Mexico determined was not qualified to testify as an expert witness in the area of firearms examinations or firearms identification. United States v. Taylor, 704 F. Supp. 2d 1192, 1195–96, 1199–1200 (D.N.M. 2009). Among her lack of qualifications was that Schwartz "has no experience in conducting firearms or toolmark identification examinations, nor has she ever taken a proficiency test in the field of firearm investigations; indeed she testified before this Court that she has never fired a gun." *Id.* at 1195.

[366] *Willock*, 596 F. Supp. 2d at 581. The judge recommended "[t]hat Sgt. Ensor not be allowed to opine that it is a 'practical impossibility' for any other firearm to have fired the cartridges other than the common 'unknown firearm' to which Sgt. Ensor attributes the cartridges." *Id.*

[367] *Id.* at 581–82. The judge also recommended that "(2)… Sgt. Ensor only be permitted to state his opinions and bases without any characterization as to degree of certainty (whether 'more likely than not' or 'to a reasonable degree of ballistic certainty'); (3) Alternatively, if you disagree with Recommendation No. 2, that Sgt. Ensor only be allowed to express his opinions 'more likely than not. . . .'" *Id.*

[368] *See generally* Medley Transcript, *supra* note 354.

[369] *Id.* at 119–20, 124–25.

[370] *Id.* at 126.

[371] *Id.*

Furthermore, the court relied heavily on a 1959 study cited by Professor Adina Schwartz from the *Willock* case,[372] but gave little weight to more recent and comprehensive firearms studies, such as the Miami-Dade study or the Ames study.[373] The judge also leaned heavily on the PCAST Report, referencing it ninety-nine times to support his decision.[374] Referring to the firearms expert's opinion testimony, the judge declared, "But at this point in time, we got to pick out what should this jury hear."[375] Then, on the record, the judge dictated exactly what the expert witness was to say on the witness stand:

> 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.[376]

At this point, the judge regressed from the "more likely than not" language he previously espoused in the *Willock* cases to only permitting an examiner to testify that such marks on a bullet were "consistent with" those on the tested firearm.[377]

---

[372] *Id.* at 89–90. The study cited by the court and Prof. Schwartz was Alfred A. Biasotti, *A Statistical Study of the Individual Characteristics of Fired Bullets*, 4 J. FORENSIC SCIS. 34, 44 (1959). *Willock*, 696 F. Supp. 2d at 559 ("A perfect correspondence between the lines on a test fired cartridge and the evidence recovered from the scene is impossible; in the real world, there is no such thing as a 'perfect match.'").

[373] *See* Medley Transcript, *supra* note 354, at 97.

[374] *Id.* at 120–22, 152.

[375] *Id.* at 14.

[376] *Id.* at 130.

[377] *Id.* at 119.

The court's decision was not about placing a so-called limitation on a forensic examiner or about the judge performing his "gatekeeper" responsibilities under Rule 702; rather, this was a judge substituting his lay opinion for that of an expert witness, instructing that expert witness to abandon his expert opinion, which identified a cartridge case from the crime scene as having been fired by the defendant's gun. The court never informed the jury of the alteration of the expert witness's testimony.[378] Meanwhile, the jury was left to determine for itself the meaning of "consistent with."[379] The court provided no scientific bases for the terminology and no definition for the jury, creating several unanswered questions: Are the marks on the questioned bullet "consistent with" just this one firearm, or are they consistent with others? If so, how many? A dozen? One hundred? How about a million other firearms? The jury has no context in which to place this testimony because the court stripped away the expert's ability to opine on the identity of the firearm that fired the questioned bullet or cartridge case. Worse, the court determined the comparison of the cartridge cases would be conducted solely by the jury without substantive guidance or assistance from any expert witness testimony.[380]

In *United States v. Davis*, another U.S. district court judge embraced the *Medley* ruling, prohibiting the examiner from opining about whether the questioned cartridge cases were a "match" to one another—or fired by a particular firearm.[381] The examiners also could not opine about whether the cartridge cases were fired by the same firearm.[382] They were only permitted to opine about whether the cartridge cases bore marks which were "consistent with" other cartridge cases or the examined firearms.[383] The judge in the *Davis* case repeatedly praised the decision of the judge in the *Medley* case,

[378] *See generally id.*

[379] *Id.* at 127–28.

[380] *Id.* at 117–18, 120–21, 127–28. "[J]uries 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." *Id.* at 118.

[381] No. 18-cr-00011, 2019 WL 4306971, at *7 (W.D. Va. Sept. 11, 2019). The court held: "Given the subjectivity of the field and the lack of any established methodology, error rate, or statistical foundation for firearm identification experts' conclusions, the testimony of the government's proposed witnesses will not be admitted in full." *Id.* at *6.

[382] *Id.* at *7.

[383] *Id.*

mentioning the judge's name no fewer than twenty-one times on the record.[384]

Other courts quickly realized the latent defects of the *Medley* and *Davis* decisions. In the 2019 case of *United States v. Johnson*, the U.S. District Court for the Southern District of New York imposed no limitations on the firearms identification expert who testified at trial.[385] The court rejected the legal reasoning in *Medley* concerning the reliability of expert testimony in the realm of firearms identification.[386] The court sharply criticized the *Medley* decision and the court's use of the term "consistent with":

> The Court also finds the resolution in *Medley*—letting "the jury itself . . . decide whether [ballistic evidence and test fires] came from the same gun,"—with no *expert* assistance—problematic. This approach invites the jury to speculate and is likely to result injury [sic] confusion. Testimony that toolmarks on casings or bullets are "consistent" with toolmarks on test-fired casings or bullets—without further explanation—provides the jury with no basis for determining whether such consistencies suggest that the ballistics evidence and test fires were fired from the same gun.[387]

A handful of other cases dot the post-PCAST landscape, each one with their own issues. In *Williams v. United States* (*Williams I*), the D.C. Court of Appeals upheld the admission of firearms identification expert testimony in which the expert testified that the bullets recovered from the crime scene matched those fired from a pistol recovered at the defendant's apartment and that the markings on those bullets were "unique."[388] The decision was in keeping with precedent too, as the D.C. appellate courts had approved of

[384] Transcript of Record, United States v. Davis, 2019 U.S. Dist. LEXIS 155037 (W.D. Va. 2019) (No. 18-cr-00011) (Urbanski, J.).

[385] (S5) 16 Cr. 281 (PGG), 2019 WL 1130258, at *19 (S.D.N.Y. 2019). Coincidentally, the *Johnson* case came from the same jurisdiction as the judge's decision in *United States v. Glynn*, 578 F. Supp. 2d 567 (S.D.N.Y. 2008)— the Southern District of New York.

[386] *Johnson*, 2019 WL 1130258, at *21 n.10.

[387] *Id.* (emphasis added) (citation omitted). In fact, the 2009 NRC Report also was sharply critical of using such loose terminology as "consistent with" or "associated with," finding there is a "problem with using imprecise reporting terminology such as 'associated with,' which is not clearly defined and which can be misunderstood." NRC REPORT, *supra* note 12, at 161.

[388] 130 A.3d 343, 347 (D.C. 2016), *aff'd on reh'g*, 210 A.3d 734 (D.C. 2019).

firearms identification expert testimony since 1923.[389] In *Williams I*, the court found that no governing precedent existed in the jurisdiction:

> There is no precedent in this jurisdiction that limits a toolmark and firearms examiner's testimony about the certainty of his pattern-matching conclusions . . . . Nor can we say that the weight of non-binding authority outside this jurisdiction is a sufficient foundation for a determination that the trial court "plainly" erred by not sua sponte limiting the toolmark examiner's testimony. We are aware of only one state supreme court decision and no federal appellate decisions limiting the opinion testimony of firearms and toolmark examiners.[390]

Yet three years later, in 2019, the D.C. Court of Appeals reversed itself, rendering an en banc opinion in the same case.[391] No new evidence had been discovered.[392] The court reversed, in part, its previous decision in *Gardner v. United States*, where the court held it was reversible error to admit the "unqualified" opinion of a firearms identification expert witness.[393] The PCAST Report, however, served as the raison d'être for the court's about-face in *Williams II*.[394] Significantly, the defendant in *Williams I* never objected to the testimony of the firearm identification expert witness, nor did the trial court judge in *Williams I* conduct an admissibility hearing regarding the testimony of the firearms identification expert.[395] This produced a trial record that contained no evidence concerning the admissibility of firearms identification expert testimony.[396] The appellate court merely took judicial

[389] *See* Laney v. United States, 294 F. 412, 416 (D.C. Cir. 1923).

[390] 130 A.3d at 347–48 (footnote omitted) (citations omitted).

[391] Williams v. United States (*Williams II*), 210 A.3d 734, 736 (D.C. 2019).

[392] *See id.* at 744.

[393] 140 A.3d 1172, 1177 (D.C. 2016). The D.C. Court of Appeals coyly refused to define what might have "qualified" the opinion of a firearms examiner in *Gardner*, stating, "We further hold that in this jurisdiction a firearms and toolmark expert may not give an unqualified opinion, or testify with absolute or 100% certainty, that based on ballistics pattern comparison matching a fatal shot was fired from one firearm, to the exclusion of all other firearms." *Id.* at 1184.

[394] *Williams II*, 210 A.3d at 741. The court would go on to hold, "[W]here the firearms and toolmark examiner not only testified, like the examiner in *Gardner*, that a specific bullet could be matched to a specific gun, but also that he did not have 'any doubt' about his conclusion. There is no question that it was error to admit this opinion testimony . . . ." *Id.* at 742.

[395] *See id.* at 738.

[396] *See id.*

notice of the 2016 PCAST Report, the 2009 NRC Report on Forensic Science, and the 2008 NRC Report on Ballistic Imaging without any other evidence in the record, then summarily pronounced the admission of such unqualified testimony was plain error.[397]

Immediately after the decision in *Williams II*, the D.C. Superior Court seized the opportunity to rule on the admissibility of firearms identification expert testimony in *United States v. Tibbs*.[398] In his eighty-five-page decision, the judge in *Tibbs* found that firearms expert testimony did not meet at least three of the five *Daubert* factors yet permitted the firearms expert to testify anyway, ordering the examiner to testify as follows:

> [T]he recovered firearm cannot be excluded as the source of the cartridge casing found on the scene of the alleged shooting—in other words, that the firearm *may* have fired the recovered casing. Mr. Coleman may not state an ultimate conclusion in stronger terms. Similarly, Mr. Coleman will be precluded at any point in his testimony from stating that individual marks are unique to a particular firearm or that observed individual characteristics can be used to "match" a firearm to a piece of ballistics evidence.[399]

The court's order had the practical effect of changing the expert's opinion from one of "identification" to one of "inconclusive."[400] The firearms

[397] *See id.* at 739–44. One wonders how the trial court's decision to admit the expert testimony of a firearms identification expert constituted "plain error" when the defendant did not object to the testimony, and caselaw in effect at the time in D.C. permitted the unqualified opinion of a firearms examiner. *See* Jones v. United States, 27 A.3d 1130, 1133 (D.C. 2011) (upholding the admission of unqualified testimony from a firearms identification expert witness). Was the D.C. Court of Appeals saying to the trial court that it committed plain error by following established precedent of the D.C. Court of Appeals?

[398] No. 2016 CF1 19431, 2019 D.C. Super. LEXIS 9, at *1, *2–3 (D.C. Super. Ct. Sept. 5, 2019).

[399] *Id.* at *77–78.

[400] AFTE defines an "*Inconclusive*" conclusion as follows:

"a. Some agreement of individual characteristics and all discernible class characteristics, **but insufficient for an identification.**

b. Agreement of all discernible class characteristics without agreement or disagreement of individual characteristics due to an absence, insufficiency, or lack of reproducibility.

c. Agreement of all discernible class characteristics and disagreement of individual characteristics, **but insufficient for an elimination**." *AFTE Range of Conclusions*, THE ASS'N OF FIREARM & TOOLMARK EXAM'RS, https://afte.org/about-us/what-is-afte/afte-range-of-conclusions (last visited Mar. 26, 2020) (emphasis added).

examiner can neither exclude the firearm nor identify it as the weapon that fired the questioned cartridge case.[401] Accordingly, under the AFTE Theory of Identification and the DOJ's Uniform Language for Testimony and Reporting, the examiner's opinion now became "inconclusive."[402] This change is surprising because the court had previously ridiculed firearms studies in which examiners returned an "inconclusive" result in black-box studies.[403] Ironically, the court determined that when the source of a bullet or cartridge case can be ascertained, but the examiner returns an "inconclusive" determination, "such responses should represent an error by the examiner."[404] It logically follows that if an examiner commits an error when he or she returns with an opinion of "inconclusive," the court has committed an error by ordering the examiner to testify that his or her opinion is tantamount to an "inconclusive" opinion.

Similar language appeared in *United States v. Shipp*.[405] There, the court found firearms identification expert witness testimony did not meet three of the five *Daubert* factors, yet the judge ruled the firearms identification expert witness was reliable enough to testify, holding that the expert witness:

> [M]ay testify that the toolmarks on the recovered bullet fragment and shell casing are consistent with having been fired from the recovered firearm, and that the recovered firearm cannot be excluded as the source of the recovered bullet fragment and shell casing. However, Detective Ring may not testify, to any degree of certainty, that the recovered firearm <u>is</u> the source of the recovered bullet fragment or the recovered shell casing.[406]

In *United States v. Adams*, the court found the testimony of the firearms examiner who testified for the prosecution did not meet most of the five

[401] *Tibbs*, 2019 D.C. Super. LEXIS 9, at *77–78.

[402] *See AFTE Theory of Identification as it Relates to Toolmarks*, *supra* note 37; DOJ FIREARMS ULTR, *supra* note 28, at 2–3; *see also* FBI ASSTR, *supra* note 26, at 2–3.

[403] *Tibbs*, 2019 D.C. Super. LEXIS 9, at *56–64.

[404] *Id.* at *60.

[405] 422 F. Supp. 3d 762, 782 (E.D.N.Y. 2019).

[406] *Id.* at 783 (emphasis in original). The judge in the *Shipp* case seemed inclined to deny the admissibility of any testimony from the firearms examiner in the case before the court, but may have felt restrained because of the Second Circuit Court of Appeals decision in *United States v. Williams*, 506 F.3d 151, 161 (2nd Cir. 2007), *cert. denied*, 552 U.S. 1224 (2008), where the court upheld the admissibility of a firearms examiner's expert testimony.

*Daubert* factors.[407] As a result, the court held the examiner in that case could only testify as to the class characteristics found in the bullet, cartridge case, and questioned firearm he examined.[408] These class characteristics included: the shape of the firing pin and the firing pin impression left on the cartridge case; the number of lands/grooves on the bullet and the barrel of the questioned firearm; whether the lands/grooves present on both were a left or right twist; plus, the caliber of the bullet, cartridge case, and the firearm itself.[409] The court permitted no discussion of individual characteristics from any of the firearms evidence, nor was the examiner allowed to provide any opinion whatsoever.[410]

In *People v. Ross*, the New York Supreme Court could not even support the terminology of "consistent with."[411] Instead, the court only permitted the firearms identification expert to testify about class characteristics and virtually nothing else.[412]

## XI. HOW FIREARMS IDENTIFICATION FULFILLS THE "DAUBERT" FACTORS FOR ADMISSIBILITY WITHOUT RESTRICTIONS OR LIMITATIONS

Before going further, an analysis of each of the five *Daubert* factors is warranted. Doing so will place the fundamental *Daubert* factors of firearms identification expert testimony into focus. These factors include: (1) whether the theory or technique has been tested (testability); (2) whether it has been subjected to publication and peer review (publication and peer review); (3) whether there is a high known or potential rate of error (error rate); (4) whether there are standards governing the technique's operation (standards); and (5) whether the theory or technique enjoys "general acceptance" within a "relevant scientific community" (general acceptance).[413]

[407] 444 F. Supp. 3d 1248, 1260–66 (D. Ore. 2020).

[408] *Id.* at 1267.

[409] *Id.*

[410] *Id.*

[411] 129 N.Y.S.3d 629, 642 (N.Y. Sup. Ct. 2020).

[412] *Id.*

[413] Kumho Tire Co. v. Carmichael, 526 U.S. 137, 149–50 (1999).

### *A. Testability*

Virtually every case regarding firearms identification expert testimony acknowledges firearms identification meets the first factor of testability.[414] Dozens of published validation studies have tested the field of firearms identification, some dating back to the time of Calvin Goddard.[415] The PCAST Report itself cites to no fewer than six of these studies in which testing was performed.[416] Even PCAST did not find firearms identification testimony was untestable.[417] To the contrary, it declared more "black box" studies should be conducted to test the discipline further.[418] In addition, hundreds of firearms examiners routinely undergo annual proficiency testing to ensure their skills of examining and identifying firearms evidence remain sharp.[419] The overwhelming number of court decisions, the PCAST Report itself, and the wealth of firearms identification studies point to but one conclusion: the field of firearms identification is testable.

### *B. Publication and Peer Review*

Even the decisions that are most critical of the firearms identification discipline have held the firearms identification discipline and the AFTE Journal satisfy the publication and peer review of *Daubert*'s second factor. In *United States v. Shipp*, the court found "the AFTE Theory has been sufficiently subjected to 'peer review and publication.'"[420] And the court in *United States v. Tibbs* disparaged the publication and peer review process of

[414] *See, e.g.*, United States v. Shipp, 422 F. Supp. 3d 762, 775–76 (E.D.N.Y. 2019).

[415] Hamby et al., *supra* note 24, at 100. Hamby's paper identifies approximately forty-three studies of the firearms and tool marks discipline from the 1930s through 2009. *Id.* at 100–04. Why none of these studies were mentioned in either of the reports from the NRC or the PCAST on Forensic Science is puzzling. The PCAST Report only reviewed empirical studies of firearms conducted since 2001. *See* PCAST REPORT, *supra* note 10, at 107–10.

[416] PCAST REPORT, *supra* note 10, at 107–10.

[417] *See id.* at 111.

[418] *Id.* at 111–13.

[419] Collaborative Testing Services (CTS) operates a forensic testing program for firearms examination. They report that approximately 630 firearms examiners took their proficiency test in firearms examination in 2020 alone. *See* COLLABORATIVE TESTING SERVS., INC,: FORENSIC TESTING PROGRAM, FIREARMS EXAMINATION TEST NO. 20-5261 SUMMARY REP. (2021), https://cts-forensics.com/reports/20-5261_Web-New.pdf; COLLABORATIVE TESTING SERVS., INC,: FORENSIC TESTING PROGRAM, FIREARMS EXAMINATION TEST NO. 20-5262 SUMMARY REP. (2021), https://cts-forensics.com/reports/20-5262_Web.pdf.

[420] 422 F. Supp. 3d 762, 777 (E.D.N.Y. 2019).

the studies published in the AFTE Journal.[421] The court found the AFTE Journal lacked double-blind peer review, its publications were unavailable to the general public, and "members of the Journal's editorial board—those who review its articles prior to publication—have a vested, career-based interest in publishing studies that validate their own field and methodologies."[422] Yet despite these observations, the court did not find that the firearms identification discipline failed the peer review and publication factor of *Daubert*.[423]

While the PCAST Report was sharply criticized for the lack of "black box" studies supporting the "foundational validity" of firearms identification expert witness testimony, it never questioned the publication and peer review of firearms identification studies in the AFTE Journal.[424] PCAST's only criticism on this topic was for the *Baldwin (Ames) Study*, which PCAST noted was subjected to neither peer review nor publication.[425] The overwhelming body of case law, the PCAST Report, and the publication of dozens of peer reviewed articles and studies by the AFTE Journal and the Journal of Forensic Science demonstrate that the firearms identification discipline meets the publication and peer review *Daubert* factor.[426]

### *C. Standards Governing the Technique's Operation*

A minority of cases have found firearms identification expert testimony does not satisfy the "standards governing the technique's operation" *Daubert* factor.[427] Much of this criticism is based on a belief that the AFTE Theory of Identification is an inadequate standard to use for identifying firearms.[428] These cases complain the AFTE Theory of Identification is "subjective and circular"[429] and that the AFTE term of "sufficient agreement" needed to make

[421] No. 2016 CF1 19431, 2019 D.C. Super. LEXIS 9, at *56 (D.C. Super. Ct. Sept. 5, 2019).

[422] *Id.* at *29–33.

[423] *See id.* at *35–36.

[424] *See* PCAST REPORT, *supra* note 10, at 104–14.

[425] *Id.* at 111.

[426] *See* United States v. Harris, 502 F. Supp. 3d 28, 40 (D.D.C. 2020) ("The Court queries whether excluding certain journals from consideration based on the type of peer review the journal employs goes beyond a court's appropriate gatekeeping function under *Daubert*.").

[427] *See, e.g.*, United States v. Shipp, 422 F. Supp. 3d 762, 782 (E.D.N.Y. 2019).

[428] *See id.* at 779–82.

[429] *Id.* at 782.

an identification "doesn't *mean* anything."[430] No person doubts the discipline of firearms identification is primarily subjective. Regardless, many courts have held "[t]he mere fact that an expert's opinion is derived from subjective methodology does not render it unreliable."[431] This is also true of other feature-comparison methods, such as the examination and analysis of latent fingerprints.[432] When considering standards pertaining to firearms identification expert testimony, a court must look at more than merely the AFTE Theory of Identification. Several other standards exist that govern the conduct of firearms examinations, the reports the examiners generate, and the way the examiners testify in court. Strangely, courts seldom discuss these other standards.

One such standard is the ULTR from the DOJ.[433] This standard governs the language, testimony, and reporting of all firearms identification experts from the DOJ, including: the Bureau of Alcohol, Tobacco & Firearms; the Drug Enforcement Agency; and the FBI.[434] It also prescribes limitations on the examiner's reports and testimony.[435] Other standards include the published internal standard operating procedures (SOPs) for every federal, state, and local forensic laboratory.[436] These SOPs govern how forensic examiners in the area of firearms identification perform their analysis of

[430] United States v. Adams, 444 F. Supp. 3d 1248, 1262 (D. Ore. 2020). Unfortunately, the analysis in the *Shipp* and *Adams* cases is misplaced. The *Daubert* factor discussed here looks for standards governing the technique's *operation*, not the technique itself. The AFTE Theory may be a technique, however it is not a standard governing its operation.

[431] *See, e.g.*, United States v. Romero-Lobato, 379 F. Supp. 3d 1111, 1120 (D. Nev. 2019). The court in *Romero-Lobato* admitted the testimony of the firearms identification expert without restriction or limitation. *Id.* at 1122.

[432] *See* United States v. Mitchell, 365 F.3d 215, 246 (3d Cir. 2004) (upholding the admission of latent fingerprint expert testimony despite the subjective nature of the examination and conclusions).

[433] DOJ FIREARMS ULTR, *supra* note 28.

[434] *Id.* at 1.

[435] *Id.*

[436] *See, e.g.*, FED. BUREAU OF INVESTIGATION, FTD TECHNICAL PROCEDURE FOR PHYSICAL AND VISUAL EXAMINATIONS (2021), https://fbilabqsd.fbi.gov/file-repository/firearms—toolmarks/operations/01-ftd-technical-procedure-for-physical-and-visual-examinations-7.pdf/view; HOUS. FORENSIC SCI. CTR., TOXICOLOGY SECTION ANALYTICAL MANUAL - STANDARD OPERATING PROCEDURES (VERSION 2.7) (2017), https://records.hfscdiscovery.org/Published/Analytical%20Manual_v2.7_Eff%202017-01-23%20to%202017-02-28.pdf; WILMINGTON POLICE DEP'T CRIME LAB'Y, FORENSIC DRUG ANALYSIS STANDARD OPERATING PROCEDURE (2015), http://www.ncids.com/forensic/labs/WPDCL/ForensicDrugSOP.pdf.

firearms evidence. At the FBI, no fewer than twenty SOPs detail how firearms examinations are conducted, and a number of these SOPs specifically govern the examination and identification of firearms, bullets, and cartridge cases.[437]

As previously mentioned, the ISO is an independent, non-governmental international organization with a membership of 164 national standards bodies.[438] It has published over 24,999 international standards on technology and manufacturing.[439] One-hundred sixty-five countries and 802 technical committees and subcommittees direct ISO in the creation of these internationally accepted scientific standards.[440] In particular, ISO/IEC 17025 is the standard that governs the calibration and testing of laboratories, including those dealing in forensics and even firearms examinations.[441] Although the 2009 NRC Report recognizes ISO standard 17025 as one of the fundamental standards governing forensic laboratories,[442] none of the courts mention it as one of the standards governing the forensic examination of firearms. Regrettably, the courts that attempted to limit or restrict the testimony of firearms identification expert witnesses never reference these standards.[443] Yet myriad standards govern the discipline of firearms identification, allowing it to pass another *Daubert* factor.

### *D. General Acceptance*

A handful of the post-PCAST court decisions have found the field of firearms identification has not achieved "general acceptance" in the relevant

[437] The FBI's firearms examination SOP and quality control documents for firearms and toolmark examinations—as well as a host of other forensic disciplines—can be found at: FBI LAB'Y QUALITY SYS. DOCUMENTS, https://fbilabqsd.fbi.gov/file-repository/firearms—toolmarks (last visited Jan. 30, 2022). The Firearms and Toolmarks Unit at the FBI has almost two dozen SOPs governing its operations and examinations. The areas governed include: "Comparison and Pattern Matching," "Bullet Examinations" "Cartridge Case Examinations" "Firearms Examinations," and "Technical Procedure for Technical and Visual Examinations."

[438] INT'L ORG. FOR STANDARDIZATION, *supra* note 266.

[439] *Id.*

[440] *Id.*

[441] INT'L ORG. FOR STANDARDIZATION, *supra* note 267267.

[442] NRC REPORT, *supra* note 12, at 113–14; *see also* INT'L ORG. FOR STANDARDIZATION, *supra* note 267.

[443] *See* United States v. Tibbs, No. 2016 CF1 19431, 2019 D.C. Super. LEXIS 9, at *65–72 (D.C. Super. Ct. Sept. 5, 2019); United States v. Shipp, 422 F. Supp. 3d 762, 779–82 (E.D.N.Y. 2019); United States v. Adams, 444 F. Supp. 3d 1248, 1266 (D. Or. 2020).

scientific or technical community. In *Adams*,[444] *Tibbs*,[445] *Shipp*,[446] and *People v. Ross*,[447] the judges believed the 2009 NRC Report and the PCAST Report serve as the talisman for determining "general acceptance" by the forensic science or technical community.[448] This reliance was badly misplaced and constitutes a legal error, because appellate courts have held the following:

> [T]he "acceptance" to which *Daubert* refers is the acceptance that the technique or theory has in the community's own field of practice when the science is being applied outside of the litigation context, not the scientific community's opinion about the standard or type of proof that should be required in litigation.[449]

Neither the 2009 NRC Report nor the PCAST Report originated from the forensic science community's field of practice. Notwithstanding this legal error, the substantive flaws of the PCAST Report have already been discussed at length in this paper, and the 2009 NRC Report's limitations with respect to using it as a metric to determine the admissibility of firearms expert

[444] 444 F. Supp. 3d at 1266–67 ("[T]hese reports suggest to me that the widespread acceptance within the law enforcement community may have created a feedback loop that has inhibited the AFTE method from being further developed. . . . Here, where the scientific community at large disavows the theory because it does not meet the parameters of science, I cannot find that the AFTE method enjoys 'general acceptance' in the scientific community.").

[445] 2019 D.C. Super. LEXIS 9, at *73–76. "The conclusions of the NRC and PCAST reports indicate that the wider academic and scientific community does not necessarily generally accept this theory. With the majority of studies published by and for the review of professional firearms and toolmark examiners, there is currently insufficient evidence that this methodology is generally accepted as proven, established, or validated—a factor that weighs against admissibility." *Id.* at *74–75.

[446] 422 F. Supp. 3d at 782–83. "For these reasons, the court finds it appropriate to consider the opinions of the authors of the NRC Report and the PCAST Report who, while admittedly not members of the forensic ballistic community, are preeminent scientists and scholars and are undoubtedly capable of assessing the validity of a metrological method. As a result, the AFTE Theory has not achieved general acceptance in the relevant community, and this factor weighs against the reliability of the AFTE Theory." *Id.* (citations omitted).

[447] 129 N.Y.S.3d 629, 639–40 (N.Y. Sup. Ct. 2020).

[448] *See Adams*, 444 F. Supp. 3d at 1166–67; *Tibbs*, 2019 D.C. Super. LEXIS 9, at *73–76; *Shipp*, 422 F. Supp. 3d at 782–83; *Ross*, 129 N.Y.S.3d at 639–40.

[449] Adams v. Lab'y. Corp. of Am., 760 F.3d 1322, 1333 (11th Cir. 2014); *see also* United States v. Baines, 573 F.3d 979, 989–91 (10th Cir. 2009) ("[T]he community is not an impartial, scientific community. . . . Consequently, while we acknowledge that acceptance by a community of unbiased experts would carry greater weight, we believe that acceptance by other experts in the field should also be considered.").

testimony are also well documented.[450] Judges would be wise to look elsewhere to gauge the "general acceptance" of the firearms identification discipline.

The courts should begin their search with accreditation. As of 2014, 88% of the nation's 409 publicly funded crime labs were accredited by an independent, professional forensic science organization.[451] Eighty-three percent of all crime labs in the United States are currently accredited by ANSI-ASQ National Accreditation Board or "ANAB."[452] ANAB bases its accreditation on ISO standard 17025 for Forensic Accreditation.[453] Accreditation can extend to twenty-five specific disciplines within the forensic science community, such as DNA (biology), latent fingerprint examinations (friction ridge analysis), toxicology, trace evidence (materials), and firearms and toolmarks examinations.[454] While only 55% (225 of 409) of all publicly funded crime labs in the U.S. conduct firearms and toolmark

[450] *See supra* Part VIII.D, regarding the 2009 NRC Report and its limitations.

[451] BUREAU OF JUST. STATS., OFF. OF JUST. PROGRAMS, U.S. DEPT. OF JUST., *supra* note 290, at 3 ("Accreditation is the process in which third-party professional forensic science accreditation bodies assess a crime lab's policies and procedures to evaluate technical competency and ability to generate valid forensic findings and interpret results. The accreditation process includes reviews of the crime lab's management practices, staff competence, training, continuing education, appropriateness of test methods, maintenance of test equipment, testing environment, handling of test items, sampling, documentation, and quality assurance of data. Professional accreditation organizations periodically monitor accredited labs to ensure crime labs maintain the standards required to remain compliant with industry best practices. Although accreditation does not guarantee that a crime lab will not make an error, it does increase confidence in the lab's ability to produce valid results by demonstrating that the lab is complying with standard operating procedures.").

[452] *Id.* at 2. ANAB was originally referred to as "ASCLD/LAB." After 2014, ANSI-ASQ National Accreditation Board or "ANAB," which acquired Forensic Quality Services International, also acquired ASCLD/LAB, International. The two accrediting organizations are now one under ANAB. *See ANAB and ASCLD/LAB Merge Forensics Operations*, ANSI NAT'L ACCREDITATION BD., https://anab.ansi.org/latest-news/anab-and-ascldlab-merge-operations (last visited Jan. 4, 2022); *see also Symposium on Forensic Expert Testimony, Daubert, and Rule 702*, 86 FORDHAM L. REV. 1463, 1483 (2018) (remarks of Dr. Alice Isenberg, speaking at the Philip D. Reed Lecture Series, Advisory Committee on Evidence Rules).

[453] *ISO/IEC 17025 Forensic Accreditation*, ANSI NAT'L ACCREDITATION BD., https://anab.ansi.org/en/forensic-accreditation/iso-iec-17025-forensic-labs (last visited August 21, 2021).

[454] *Directory of Accredited Organizations*, ANSI NAT'L ACCREDITATION BD., https://search.anab.org/ (last visited April 6, 2020).

examinations,[455] a search of the ANAB website reveals that ANAB recognizes that 251 crime labs are accredited in the field of firearms and toolmark testing and inspection.[456] The list of crime labs in the United States accredited for their firearms and toolmarks identification expertise include virtually every state in the nation and cities such as New York, Los Angeles, Boston, Washington, Denver, Houston, Miami, Philadelphia, Phoenix, St. Louis, Dallas, and Austin, plus the crime labs for the Department of Defense, DEA, ATF, and the FBI.[457] This is what "general acceptance" looks like.

ANAB accreditation entails more than just adherence to ISO standard 17025. It also requires training of examiners, testimony monitoring, validation of procedures, and annual proficiency testing.[458] To maintain laboratory accreditation, firearms identification expert witnesses must undergo annual proficiency testing to determine whether the examiners perform to industry standards.[459] In the United States, 98% of all crime labs conduct proficiency testing of their examiners,[460] and the dominant company performing these tests on behalf of the crime labs is Collaborative Testing

[455] BUREAU OF JUST. STATS., OFF. OF JUST. PROGRAMS, U.S. DEPT. OF JUST., NCJ 250151, PUBLICLY FUNDED FORENSIC CRIME LABORATORIES: RESOURCES AND SERVICES, 2014, at 2 (2016).

[456] *Directory of Accredited Organizations*, ANSI NAT'L ACCREDITATION BD., http://search.anab.org/?__hstc=4076783.cbdc1b4f8f70a858a2758acd67b0bed0.1585746118799.1585746118800.1585746118800.1&__hssc=4076783.1.1585746118800&__hsfp=3978889890 (last visited April 6, 2020). The number of accredited crime labs in the area of firearms testing is larger than the number of crime labs offering such testing because ANAB also accredits crime labs from outside the United States. *ISO/IEC 17025 Forensic Accreditation*, *supra* note 453.

[457] *Directory of Accredited Organizations*, *supra* note 456.

[458] *See* BUREAU OF JUST. STATS., OFF. OF JUST. PROGRAMS, U.S. DEPT. OF JUST., *supra* note 290, at 3. According to the DOJ, proficiency testing is defined as "a quality control tool used to examine the performance of the crime lab personnel and to determine whether personnel are following industry standards. To receive and maintain professional accreditation, a crime lab is required to evaluate the technical competence of analysts, other personnel, and the overall performance of the crime lab through proficiency testing. Proficiency tests are conducted internally or externally using declared tests (an examiner knows the sample to be analyzed is a test sample), random case reanalysis (an examiner's work is randomly selected for reanalysis by another examiner), and blind tests (the examiner or crime lab is not aware of being tested)." *Id.*

[459] *See id.*; *see also ISO/IEC 17043 PT Providers: Accredited Forensic Test Providers*, ANSI NAT'L ACCREDITATION BD., https://anab.ansi.org/en/forensic-accreditation/proficiency-testing (last visited August 21, 2021) (listing all forensic proficiency test providers who are accredited to ISO/IEC 17043 (Conformity Assessment - General Requirements For Proficiency Testing) and are a required part of ANAB accreditation).

[460] BUREAU OF JUST. STATS., OFF. OF JUST. PROGRAMS, U.S. DEPT. OF JUST., *supra* note 290, at 1.

Services (CTS).[461] CTS conducts proficiency testing in ten different forensic disciplines, including DNA, toxicology, latent fingerprints, anthropology, and firearms & toolmarks. Its proficiency tests extend to over 900 crime labs in eighty different countries worldwide.[462] CTS's proficiency tests are accredited to ISO/IEC standard 17043.[463] Twice a year, CTS provides a proficiency test specifically for "firearm examination," which requires an examiner to identify which bullet or cartridge case was fired from a particular firearm.[464] Four questioned bullets or cartridge cases are provided to the examinee for testing along with three known bullets or cartridge cases.[465] Depending on the examination, at least one or more of the four questioned bullets or cartridge cases are a "match."[466] In 2018 and 2019, CTS conducted a total of 1,191 proficiency tests in firearms examination, with 1,172 respondents returning the correct conclusion.[467] This correlates to an error rate of less than 1.4%.[468] The widespread use of proficiency tests in firearms examination and the correspondingly low error rate further demonstrates the "general acceptance" of the firearms identification discipline.

Next, the National Integrated Ballistics Information Network (NIBIN) is a "national network that allows for the capture and comparison of ballistic evidence to aid in solving and preventing violent crimes involving firearms."[469] Since its inception in 1997 by the Bureau of Alcohol, Tobacco and Firearms (ATF), NIBIN has acquired over 16 million images of bullets, cartridge cases, and other ballistic data from over 4.5 million pieces of

[461] *Collaborative Testing Experts Is Your Proficiency Testing Expert*, COLLABORATIVE TESTING SERVS.: FORENSICS TESTING PROGRAM, https://cts-forensics.com/index-forensics-testing.php#row-2 (last visited August 21, 2021).

[462] *Id.* (under "CTS Forensics History").

[463] *Id.* (under "Who is CTS?").

[464] *See Firearms & Toolmarks*, COLLABORATIVE TESTING SERVS.: FORENSICS TESTING PROGRAM, https://cts-forensics.com/program-3.php. The schedule and summary reports tabs link to the twice annual schedule of proficiency testing for firearms examination.

[465] *See* COLLABORATIVE TESTING SERVICES, INC., FIREARMS EXAMINATION TEST NO. 20-5262 SUMMARY REPORT 3 (2020), https://cts-forensics.com/reports/20-5262_Web.pdf.

[466] *Id.*

[467] *See Firearms & Toolmarks*, *supra* note 464 (results for CTS's proficiency test in firearms identification).

[468] This error rate is somewhat analogous to the error rate reported in the six post-NRC report firearms identification studies from 2009-2016. *See supra* Part VI.

[469] *Fact Sheet - National Integrated Ballistic Information Network*, BUREAU OF ALCOHOL, TOBACCO, FIREARMS & EXPLOSIVES (Sept. 2021), https://www.atf.gov/resource-center/fact-sheet/fact-sheet-national-integrated-ballistic-information-network (last visited April 6, 2020).

evidence, resulting in over 132,000 confirmed investigative leads or "hits" from the NIBIN database over the past two decades.[470] While NIBIN cannot *identify* a particular firearm (only a trained examiner can), it operates under the core assumption that firearms produce individual characteristics and are therefore *identifiable*.[471] The success of NIBIN, utilized by over 300 law enforcement agencies nationwide, is another indication of the "general acceptance" of firearms identification.

Additionally, the U.S. Department of Commerce's NIST recognizes the general principles of firearms identification and has created the OSAC to develop and publish national standards in all fields of forensic science.[472] The NIST's sponsorship of the OSAC includes the Firearms & Toolmarks subcommittee, consisting of nineteen members and three officers from the firearms and toolmark forensic community.[473] Then there remains the 1964 report from the Warren Commission regarding the assassination of President John F. Kennedy. Commissioned by Congress and chaired by the Chief Justice of the U.S. Supreme Court, the Warren Commission relied on and accepted firearm identification expert witness testimony to determine who assassinated the President of the United States.[474] In doing so, all three branches of the U.S. government relied on firearms identification to answer the question of who shot the President.

Accreditation, widespread proficiency testing, the success of ATF's NIBIN database, the Commerce Department's recognition of firearms identification, and the reliance of the U.S. government on firearms identification to investigate and solve the assassination of a U.S. president serve as cornerstones for the "general acceptance" of the firearms identification discipline that district courts overlooked.

---

[470] *Id.* NIBIN reports, "As of February 2019, the NNCTC has conducted 127,917 correlation reviews, resulting in the generation of more than 33,000 investigative leads to law enforcement partners. These leads help solve homicides, attempted homicides, robberies, and other non-fatal shooting incidents." *NIBIN National Correlation and Training Center*, BUREAU OF ALCOHOL, TOBACCO, FIREARMS & EXPLOSIVES (Sept. 21, 2021), https://www.atf.gov/firearms/nibin-national-correlation-and-training-center.

[471] "However, it is important to note that the final determination of a match is always done through direct physical comparison of the evidence by a firearms examiner, not the computer analysis of images." NRC REPORT, *supra* note 12, at 153.

[472] *See Firearms & Toolmarks Subcommittee*, NAT'L INST. OF STANDARDS & TECH. (NIST) (Jan. 27, 2022), https://www.nist.gov/osac/firearms-toolmarks-subcommittee (last visited Apr. 6, 2020).

[473] *Id.*

[474] WARREN COMMISSION REPORT, *supra* note 9, at 58–66.

## *E. Error Rate*

When looking at error rates, the question is "[w]hether, in respect to a particular technique, there is a high 'known or potential rate of error.'"[475] The U.S. Supreme Court was justifiably concerned about a "high" rate of error; however, it did not establish what constituted a "high" rate of error.[476] In each of the post-PCAST Report cases that restricted firearms expert witnesses' testimony, the error rate issue was the lynchpin of their legal analysis. In *United States v. Adams*, the court blindly accepted the PCAST Report's error rates for the studies it cited—as high as 2.2%—without verifying the accuracy of the error rate.[477] In *United States v. Shipp*, the court also accepted the PCAST Report's questionable claim that the error rates in the Miami-Dade and Baldwin (Ames) studies were 2.1% and 2.2%, respectively.[478] Again, the court did not verify those claims by examining the studies themselves.[479] The judge in *United States v. Davis* failed to discuss error rate altogether,[480] and the court in *United States v. Tibbs* attacked the studies PCAST Report cited, while ignoring the inconveniently low error rates reported by those studies.[481] The variously calculated error rates for "identification" decisions by firearms examiners in the six studies cited by PCAST can be summarized as follows:[482]

*Hamby Study*= 0%[483]

*Fadul Study*= 0.063%[484]

[475] Kumho Tire Co. v. Carmichael, 526 U.S. 137, 149 (1999).

[476] *See id.*

[477] 444 F. Supp. 3d 1248, 1264–65 (D. Or. 2020).

[478] 422 F. Supp. 3d 762, 777–78 (E.D.N.Y. 2019).

[479] *See id.* at 777–79.

[480] *See* No. 4:18-cr-00011, 2019 U.S. Dist. LEXIS 122135, at *1 (W.D. Va. July 23, 2019).

[481] *See* No. 2016 CF1 19431, 2019 D.C. Super. LEXIS 9, at *36, *65 (D.C. Super. Ct. Sept. 5, 2019). The Court in *Tibbs* wrote an eighty-five-page opinion, devoting thirty pages of the ruling to a discussion on the studies that the PCAST Report discussed, yet the court never mentioned a single error rate from any of these studies.

[482] Given that the PCAST Report was not peer reviewed and erroneously misrepresented the error rates for the Miami-Dade and Baldwin (Ames) studies, it is more accurate to quote the error rates reported by the studies themselves, not those reported in the PCAST Report.

[483] *Hamby Study*, *supra* note 24, at 107.

[484] *Fadul Study*, *supra* note 484192, at 385.

*Stroman Study*= 0%[485]

*Miami-Dade Study*= 0.7%[486]

*Baldwin (Ames) Study*=1.01%[487]

*Smith Study*= 0.14%[488]

While these six studies demonstrate a remarkably low error rate for the firearms identification discipline, the studies only provide estimates of *raw* error rates. The studies do not consider the widespread use of quality assurance and control measures that most forensic laboratories utilize, such as independent verification and peer review, which serve to further mitigate any potential error by the individual examiners.[489] The majority of these studies also incorporate the examination of bullets and cartridge cases fired from consecutively manufactured barrels or slides where sub-class characteristics are present and which could potentially mislead the examiner.[490] This particular design factor in these studies makes firearms identification studies *more* difficult than the vast majority of casework from real-life shootings, where criminals do not use multiple firearms equipped with consecutively made barrels or slides, nor the same caliber of firearms,

[485] *Stroman Study*, *supra* note 206, at 159.

[486] *Miami-Dade Study*, *supra* note 184, at 29.

[487] *Baldwin Study*, *supra* note 211, at 3. As part of the *Baldwin (Ames) Study*, test subjects were required to not use any of their laboratories' quality control measures, including peer review and verification. *Id.* at 6. The *Miami-Dade Study* also did not permit respondents to utilize technical review. *Miami-Dade Study*, *supra* note 184, at 39.

[488] *Smith study*, *supra* note 221, at 943.

[489] *See id.* at 945. *See also Baldwin Study*, *supra* note 211, at 5, 18 (where participants were required to not use any peer review process or verification in reaching their conclusions). When an examiner is participating in firearms studies they are removed from the laboratory's QA system and are asked to evaluate the samples independently. Their training, experience, and application of the AFTE theory when making a decision is what is being studied.

[490] *Miami-Dade Study*, *supra* note 184, at 37 ("Consecutively manufactured barrels represent the best possibility for the production of two firearms that could produce non-distinguishable markings since the same tools and machining processes are utilized back-to-back on one barrel after another. This process thus represents a situation where the most similarity should be seen between barrels. If there were ever any chance for duplication of individual marks, it would occur during the manufacture of consecutively manufactured barrels . . . . Once the specter of subclass influence is eliminated, each firearm/tool produces a signature of identification (striation/impression) that is unique to that firearm/tool.").

or even firearms from the same manufacturer.[491] Given these circumstances, it seems likely the false positive error rates in the firearms identification field are significantly *lower* than those reported in many of these empirical studies. Regardless, these firearm identification studies consistently demonstrate that when a qualified firearms examiner opines they have "identified" a bullet or cartridge case as having been fired from a particular firearm, the firearm examiner's conclusion is accurate approximately 99% of the time. No less than three post-PCAST "black box" firearms identification studies confirm this accuracy.

In 2018, Mark Keisler led a team of forensic examiners from the Indiana State Police Crime Laboratory in a firearms study comparing cartridge cases fired from nine .40 S&W semi-automatic pistols manufactured by HK, Smith & Wesson, and Glock.[492] Examiners were each given a kit that contained twenty individual sample envelopes, which contained two cartridge cases each.[493] Keisler's team instructed the participants to only compare the two cartridge cases in each separate sample envelope and reach.[494] This constituted a "black box"[495] study, just as PCAST demanded; however 126 firearms examiners who participated in the study accurately recorded 1,508 identifications and 805 exclusions. [496] None of the participants recorded a single false identification or false exclusion.[497] This equates to an error rate of 0%.[498]

In 2020, Jaimie Smith of the Prince Georges County Police Department in Landover, Maryland, published a peer-reviewed study in the Journal of

[491] One exception to this might be an officer-involved shooting where two or more law enforcement officers discharge their firearms at a crime scene. In such a case, their firearms may be of the same make, caliber, and manufacturer. They might even have consecutively manufactured barrels; however this would be a very rare occurrence.

[492] Mark A. Keisler et al., *Isolated Pairs Research Study*, 50 AFTE J. 56, 56 (2018).

[493] *Id.* at 57.

[494] *Id.*

[495] *See id.* at 57–58 (In this particular study, "[p]articipants were instructed to only compare the two cartridge cases in each individual sample envelope and come to a conclusion of Identification, Exclusion, or Inconclusive. . . . A participant's answer to one set was not dependent on the answers from another set, which has been a criticism of some types of research. As evident by the results, 126 participant examiners were able to reach correct conclusions with no false positives or false negatives.").

[496] *Id.* at 57.

[497] *Id.*

[498] *Id.*

Forensic Sciences,[499] which utilized the comparison of bullets fired from thirty consecutively manufactured Berretta 9mm semi-automatic pistol barrels.[500] Smith designed the study to conform to the model recommended by PCAST report, including a "black box, open set" study, conducted by a third-party with no stake in the outcome of the study.[501] Seventy-four examiners participated in this firearms identification study, requiring each of them to compare up to sixty-five different bullets, then record their conclusions.[502] The results found examiners correctly identified the source of a bullet 1,257 times and falsely identified a bullet just once.[503] With regards to eliminations, the participants correctly eliminated bullets 10,935 times and falsely eliminated bullets only eighteen times.[504] This computes to an overall error rate of 0.16%.[505] Inconclusive results were not counted as errors.

Finally, in 2021, the FBI released perhaps the most comprehensive "black box" study on firearm identification to date.[506] This firearms study recruited 173 firearms examiners who made over 20,000 comparisons of both cartridge cases and bullets from consecutively manufactured slides and barrels.[507] The study answered the PCAST Report's demand for an additional "black box, open set" study of the firearms identification discipline, employing the U.S. Department of Energy's Ames Laboratory to conduct the

[499] Jaimie A. Smith, *Beretta Barrel Fired Bullet Validation Study*, 66 J. FORENSIC SCI. 547, 547 (2021).

[500] *Id.* at 548.

[501] *Id.* Smith employed CTS to perform the study.

[502] *See id.* at 548–51.

[503] *Id.* at 552.

[504] *Id.* at 551.

[505] *Id.* While inconclusive results were not counted as errors, if they had been so counted, the overall error rate would still have been only 52 of 1303 or 4.0% of submitted results. *Id.* at 554.

[506] L. SCOTT CHUMBLEY ET AL., ACCURACY, REPEATABILITY, AND REPRODUCIBILITY OF FIREARM COMPARISONS 3 (2021) [hereinafter 2021 FBI FIREARMS STUDY].

[507] *Id.* at 1, 4. This study entailed 173 firearms examiners who made over 20,000 comparisons of both cartridge cases and bullets from consecutively manufactured slides and barrels. The results of this study confirm the low rates of error reported in prior studies.

research.[508] The FBI Study reported a false positive error rate 0.656% for bullets and 0.933% for cartridge cases, respectively.[509]

The PCAST Report itself notes that a false-positive rate lower than 5% is the benchmark for reliability.[510] The issue for PCAST, however, was not the error rates reported by the six firearms studies it surveyed; it was the way those studies were conducted. PCAST believed "closed set" studies, where the answer was always present in the collection provided to the examiner, substantially underestimated the error rate of the firearms identification discipline.[511] Instead, the PCAST Report preferred "open set" studies where the answer or solution may not have been present in the collection given to the examiner.[512] The *Miami-Dade Study* and the *Baldwin (Ames) Study* were typical of this "open set" design, and examiners partaking in these studies did not know whether the matching cartridge case or bullet was part of the collection submitted for their examination. PCAST claimed the false-positive rates in "open set" studies were 100-fold higher than in the "closed set" studies.[513]

While this sounds dramatic, in reality some of the firearms studies reported a near-zero percent false-positive rate, and the highest suggested false-positive rate (the *Miami-Dade Study*) came in at just 1.2%.[514] While significant, the minor differences in calculated error rates still demonstrate a consistently low error rate of false identifications for the firearms

[508] PCAST REPORT, *supra* note 10, at 113. Dr. Eric Lander, then co-chair of PCAST, stated the following in 2018: "With only a single well-designed study estimating accuracy, PCAST judged that firearms analysis fell just short of the criteria for scientific validity, which requires reproducibility. A second study would solve this problem." Eric S. Lander, *Fixing Rule 702: The PCAST Report and Steps to Ensure the Reliability of Forensic Feature-Comparison Methods in the Criminal Courts*, 86 FORDHAM L. REV. 1661, 1672 (2018). That second study has now arrived. Even by PCAST's standards, firearms identification testimony must now be recognized as having achieved "scientific validity."

[509] 2021 FBI FIREARMS STUDY, *supra* note 506, at 16.

[510] PCAST REPORT, *supra* note 10, at 152 ("To be considered reliable, the FPR [false positive rate] should certainly be less than 5 percent and it may be appropriate that it be considerably lower, depending on the intended application.").

[511] *Id.* at 111.

[512] *Id.* at 106–08.

[513] *Id.* at 109.

[514] *Miami-Dade Study*, *supra* note 184, at 33. The actual report had a calculated error rate of 0.7% with a 95% confidence interval of 1.2%. PCAST reported the error rate at 4.7% at the 95% confidence interval. PCAST REPORT, *supra* note 10, at 111. What was the difference? PCAST calculated false positives among conclusive examinations.

identification discipline. Despite what PCAST claims, the study design does not have any meaningful impact on the error rate results. It simply does not matter whether a study is "black box," "white box," "closed," or "open." All firearms studies return a consistently low false-positive error rate. Even the widespread proficiency testing of the discipline returned an error rate of only 1.4%, further supporting the reliability of the firearms identification discipline. The PCAST Report also confines itself to a belief that "black box" studies constitute the *sole* means to assess the scientific validity of a forensic discipline, a proposition the American Society of Crime Laboratory Directors rejected.[515]

So, what is an acceptable error rate for firearms identification? First, one must recall "error rate" is just one of five *Daubert* factors. The U.S. Supreme Court held that those factors do not constitute "a definitive checklist or test" for admissibility.[516] "*Daubert* itself is not to the contrary. It made clear that its list of factors was meant to be helpful, not definitive."[517] Second, the U.S. Supreme Court was concerned about *high* error rates.[518] The existence of a *low* or no-error rate has never been an impediment to the admissibility of feature comparison or pattern-based evidence, such as firearms identification or latent fingerprint comparison.[519] Yet two of the courts have found virtually

[515] AM. SOC'Y OF CRIME LAB'Y DIR., INC., STATEMENT ON SEPTEMBER 20, 2016 PCAST REPORT ON FORENSIC SCIENCE 1 (2016) [hereinafter ASCLAD] ("PCAST's dismissal, however, of a wealth of existing research because it does not meet an arbitrary criteria of black box studies with an ideal sample size is unhelpful. . . . ASCLD disagrees with discarding these studies as not credible simply for lack of black box studies or ideal sample size. ASCLD concurs that black box and white box studies are significantly important and helpful. . . . ASCLD does not agree, however, that black box studies are the singular method through which to judge an entire forensic discipline's reliability."). ASCLAD has merged with "ANAB" which is the primary independent organization to accredit forensic laboratories.

[516] Daubert v. Merrell Dow Pharms., Inc., 509 U.S. 579, 593 (1993).

[517] Kumho Tire Co. v. Carmichael, 526 U.S. 137, 151 (1999).

[518] *Id.* at 149.

[519] *See, e.g.*, United States v. Straker, 800 F.3d 570, 631–32 (D.C. Cir. 2015) (upholding the admission of expert testimony on the identification of latent fingerprints despite the expert being unable to proffer a known human error rate in the discipline and a 0% rate of error for the ACE-V methodology); United States v. Baines, 573 F.3d 979, 983–84, 988–89, 991 (10th Cir. 2009) (unable to discern the human error rate for latent fingerprints, yet admitting the testimony of a latent print examiner who identified fingerprints in the case as being made by the defendant); United States v. Mitchell, 365 F.3d 215, 239–41, 246 (3rd Cir. 2004) (upholding the admissibility of latent fingerprint identification testimony without any specific error rate(s)); United States v. Herrera, 704 F.3d 480, 483–87 (7th Cir. 2013) (upholding admission of a latent fingerprint identification expert witness with no definitive error rate); United States v. Gutierrez-Castro, 805 F. Supp. 2d 1218,

any error rate for the firearms identification field as too high. A closer examination of these cases is in order.

In *United States v. Adams*, the judge opined that firearms identification having an error rate of 2.2% could lead to a wrongful conviction in one out of forty-six cases.[520] The court conceded that, "Even at its worst, comparison analysis has a very low rate of error."[521] Ironically, the judge then determined that the error rate for firearms identification was "far too high."[522] Regardless, the court's faulty logic assumes every error in a firearms identification conclusion translates into a wrongful conviction.[523] This is simply untrue. A forensic lab's quality control and peer-review process, other evidence in the case, cross examination by the defense counsel, defense examination of the ballistic evidence, possible defense expert witness testimony regarding the same, jury instructions, the prosecution's burden of proof beyond a reasonable doubt, and even scrutiny of the expert's testimony by the jury itself, all act as safeguards against such an outcome.

The judge in *United States v. Shipp* also found that PCAST's reported error rate of 2.2% made the AFTE Theory and the firearms identification discipline unreliable.[524] Then, by way of comparison, the court held up the error rate in DNA as a forensic discipline with an acceptable rate of error of one in ten billion.[525] If this constitutes the new benchmark for admissibility, then the federal courts have crossed the Rubicon and rewritten Rule 702, no longer requiring an expert witness's opinion to be reliable, but mandating it

1232–34 (D.N.M. 2011) (upholding the admission of a latent print expert with only an unquantified "low" rate of error in conducting identification by fingerprints).

[520] 444 F. Supp. 3d 1248, 1264 (D. Or. 2020).

[521] *Id.* at 1266.

[522] *Id.* at 1265. Besides the "wrongful conviction" fallacy, the *Adams* decision committed other errors. It blindly accepted PCAST's incorrect error rates instead of verifying the actual, lower error rates reported by the studies themselves. The judge also declared the Baldwin (Ames) firearms study was a "closed set" and suggested the error rate of 2.2% could be even higher had it been an "open set." *Id.* at 1264–65. In fact, the PCAST Report acknowledged the *Baldwin (Ames) Study* was actually an "open set" with the highest reported error rate. PCAST REPORT, *supra* note 10, at 110–11.

[523] *See Adams*, 444 F. Supp. 3d at 1264–65.

[524] 422 F. Supp. 3d 762, 777–79 (E.D.N.Y. 2019).

[525] *Id.* at 778–79.

be nearly perfect or flawless.[526] Yet, as the Seventh Circuit has held, "[e]vidence doesn't have to be infallible to be probative."[527]

Furthermore, the 2.2% error rate cited by these courts is simply wrong. All the studies cited in the PCAST report delivered a false-positive rate no greater than 1.01%.[528] The PCAST report claimed higher error rates by throwing out the inconclusive results and reporting only the upper bound of the 95% confidence limit to maximize the potential false positive error rate.[529] The trial courts seemed unaware of this distortion by PCAST.

The third trial court to severely limit firearms identification expert testimony, in *United States v. Tibbs*,[530] took issue with the studies that reported the error rate; however, the court avoided any discussion of those error rates, focusing instead on the court's belief that the study's design methodology and peer review were too flawed to confer legitimacy on the reported error rates.[531] Here, the court committed an unforced legal error, for the Eleventh and Ninth Circuits have both held: "In most cases, objections to the inadequacies of a study are more appropriately considered an objection going to the weight of the evidence rather than its admissibility."[532]

The *Tibbs* court did cite two studies whose "open set" design it found acceptable: the *Baldwin (Ames) Study* from 2014 and the post-PCAST Report firearms study conducted by Mark Keisler.[533] The court conceded these two studies returned false-positive rates of 1.01% and 0%, respectively, but the

[526] Unfortunately, perfection does seem to be the emerging standard courts are using to weigh the admissibility of expert testimony. Recently, a U.S. District Court judge found even expert testimony concerning DNA evidence should not be admitted because it is not flawless. *See* United States v. Gissantaner, 417 F. Supp. 3d 857, 886 (W.D. Mich. 2019), *rev'd*, 990 F.3d 457 (6th Cir. 2021) ("The DNA evidence sought to be admitted in this case—in essence, that it is 49 million times more likely if Daniel Gissantaner is a contributor to the DNA on the gun than if he is not—is not really evidence at all. It is a combination of forensic DNA techniques, mathematical theory, statistical methods (including Monte Carlo-Markov Chain modeling, as in the Monte Carlo gambling venue), decisional theory, computer algorithms, interpretation, and subjective opinions that cannot in the circumstances of this case be said to be a reliable sum of its parts. Our system of justice requires more.").

[527] United States v. Herrera, 704 F.3d 480, 486 (7th Cir. 2013).

[528] PCAST REPORT, *supra* note 10, at 108.

[529] *Id.* at 110–11.

[530] No. 2016 CF1 19431, 2019 D.C. Super. LEXIS 9, at *80–81 (D.C. Super. Ct. Sept. 5, 2019).

[531] *See id.* at *48, *55–56.

[532] Adams v. Lab. Corp. of Am., 760 F.3d 1322, 1334 (11th Cir. 2014); Hemmings v. Tidyman's Inc., 285 F.3d 1174, 1188 (9th Cir. 2002).

[533] 2019 D.C. Super. LEXIS 9, at *30, *47–48, *56.

court ignored even these findings because they were neither peer reviewed nor published in another publication besides the AFTE Journal.[534] The court avoided the fact that the Keisler study was yet another data point, again confirming the consistently low error rates of the firearms identification discipline.

The mental gymnastics employed by these courts to avoid discussing the error rates of firearms identification experts continues with the "inconclusive fallacy." In real life and in many of the firearms identification studies, examiners are allowed to provide an "inconclusive" response if they were unable to either identify or exclude a bullet or cartridge case as having been fired by a particular firearm.[535] PCAST focused heavily on the inconclusive rate of some of the studies cited in its report, noting a disparity between the inconclusive rates for "open set" firearms studies versus "closed set" studies.[536] Because of the rate or number of inconclusive conclusions that firearms examiners reached in these studies, PCAST and at least a handful of courts seem fixated on them. In *United States v. Shipp*, the judge found the "inconclusive" rates of these studies cast doubt on the accuracy of the error rates or false-positive rates reported in other firearms studies.[537] The judge in *United States v. Tibbs* came to a similar conclusion, pulling the inconclusive rate for firearms identification into the court's calculus of an error rate.[538]

> Focusing on the only two "open" studies, the Ames Laboratory study calculated a false positive error rate of 1.01%, while the Keisler study reported a false positive error rate of 0%. If the inconclusives are considered as errors, however, the Ames Laboratory study's error rate among different source comparisons soars to 34.76% while the Keisler study's error rate rises to 20.14%. Again, Dr. Scurich's approach of treating inconclusives as false

[534] *Id.* at *63–64.

[535] DOJ FIREARMS ULTR, *supra* note 28, at 2–3 ("The basis for an 'inconclusive' conclusion is an examiner's opinion that there is an insufficient quality and/or quantity of individual characteristics to identify or exclude."). *See also* WARREN COMMISSION REPORT, *supra* note 9, at 85 ("Under such circumstances . . . there are not sufficient individual characteristics to enable the examiner to make a firm identification."); *id.* at 553 ("A bullet or cartridge case cannot always be identified with the weapon in which it was fired.").

[536] PCAST REPORT, *supra* note 10, at 109.

[537] 422 F. Supp. 3d 762, 784 n.10 (E.D.N.Y. 2019).

[538] *See* No. 2016 CF1 19431, 2019 D.C. Super. LEXIS 9, at *57–64 (D.C. Super. Ct. Sept. 5, 2019).

> positives does not appropriately address the issue presented by inconclusives, but the large number of the inconclusives reported in the studies greatly reduces their persuasive force in establishing the ability of a firearms and toolmark examiner to make accurate source determinations.[539]

Appellate courts have spurned this reasoning. If the trial court were looking at firearms studies with relatively high inconclusive rates and high false positive rates, a judge might be on firm ground to reach that conclusion. However, where judges have confronted *high* inconclusive rates and *low* false-positive rates, the courts have ruled that this is a benchmark of reliability.[540] Under these circumstances, the keystone is the false-positive rate.[541] In fact, the U.S. Court of Appeals for the Third Circuit addressed this exact situation with another feature-comparison discipline—latent fingerprints. In *United States v. Mitchell*, the appellate court found the following:

> While a system of identification with a high false negative rate may be unsatisfactory as a matter of law enforcement policy, in the courtroom the rate of false negatives is immaterial to the *Daubert* admissibility . . . offered to prove positive identification because it is not probative of the reliability of the testimony *for the purpose for which it is offered* (i.e., for its ability to effect a positive identification). Thus we must focus on evidence that is probative of the rate of false positives . . . [A] method may be designed to lower its false positive error rate by accepting a large number of false negatives out of an abundance of caution.[542]

---

[539] *Id.* at 62–63.

[540] *See, e.g.*, United States v. Mitchell, 365 F.3d 215, 239–40 (3d Cir. 2004).

[541] *See id.*

[542] *Id.* at 239–40, n.19. To paraphrase the Third Circuit in note 19 of its opinion in *Mitchell*: While a significant "inconclusive" rate might suggest a generally error-prone method, it is equally consistent with a very conservative method with a low false positive error rate. That is, a method may be designed to lower its false positive error rate by accepting a large number of false negatives out of an abundance of caution. *See id.* One very familiar example of such a system is the criminal jury using the "beyond a reasonable doubt" standard: As the adage (attributed to Blackstone) says, "It is better that ten guilty escape [inconclusives] than one innocent suffer [a false positive]." The same may be true for firearms identification—the examiners who declared they could not match a firearm in the Miami-Dade and Baldwin (Ames) firearms studies (the examiners responsible for the

The inconclusive rates reported in some firearms studies do not support the court's reasoning in *Tibbs*; indeed, the inconclusive rates undermine the court's logic because the low false-positive rate for firearms identification is a function of the significant "inconclusive" rates returned in these studies. By giving examiners the ability to decline an identification conclusion of which they are uncertain, the firearms discipline maintains the integrity to consistently identify bullets or cartridge cases with remarkable accuracy, thereby reducing the number of false positives. This process ensures an examiner's conclusion as to identification is reliable. Given the context of a consistently low false-positive rate, any discussion of a high inconclusive rate in the firearms identification discipline is "full of sound and fury, / [s]ignifying nothing."[543]

## XII. ABUSE OF DISCRETION: HOW THE COURTS GOT IT WRONG ON FIREARMS EXPERT TESTIMONY

After all the long opinions dissecting the admissibility of firearms identification evidence in *Adams*, *Tibbs*, *Shipp*, *Davis*, *Williams*, and *A.M.*, none of these cases held that such testimony was inadmissible or denied the ability of a firearms expert witness to testify at trial. By doing so, the courts acknowledged such testimony was reliable, lest it run afoul of Rule 702. Yet the courts committed another abuse of discretion in dealing with firearms identification expert testimony. The abuse of discretion can be cataloged four ways. First, the courts dictated "limitations" to the expert's testimony, which had the practical effect of changing the witnesses' opinion and substituting the judge's opinion for the expert's conclusions. Second, the language the court demanded of the experts was based in neither law nor science. Third, the courts conflated their role as "gatekeeper" with the trier of fact. Fourth, the courts denied the jury the opportunity to hear an expert's true opinion regarding the identification of a firearm, despite its proven reliability. Any of

putative inconclusive results) may have done so because they would rather commit to an inconclusive result rather than risk a small chance of a false positive identification.

[543] WILLIAM SHAKESPEARE, MACBETH act 5, sc. 5, 2289–90. In actual casework, a significant number of comparisons of known and questioned bullets and/or cartridge cases end in an "inconclusive" result. In the Warren Commission report, the FBI firearms examiners were unable to find a .38 caliber revolver carried by Lee Harvey Oswald, was the firearm used to murder Dallas Police Officer J.D. Tibbet, despite having Oswald's .38 Smith & Wesson revolver and recovering four intact .38 caliber bullets from Tibbet's body. WARREN COMMISSION REPORT, *supra* note 9, at 558–60.

these failures by the trial courts would constitute an abuse of discretion, let alone a combination of them.

### *A. "Limitations" are Used as a Pretext to Rewrite a Witness' Testimony*

Federal Rule of Evidence 702 and the comments that accompany it provide the guidance judges need to determine the admissibility of expert witness testimony. Chief among these is the concept of reliability and whether the expert's testimony is the product of reliable principles and methods, and that the expert has reliably applied those principles and methods to the facts of the case.[544] Judges serve as the "gatekeeper" for the admissibility of expert witness testimony and enjoy considerable leeway in determining whether a particular expert witness in a specific case is reliable. Such a determination by the trial court judge and his or her decision to admit or exclude expert witness testimony will only be overturned by an appellate court for an "abuse of discretion."[545]

However, when a judge finds an expert's principles or methods unreliable, the trial court's sole course of action is to deny the admissibility of the expert's testimony. Federal Rule of Evidence 702 and the Comments to the rule, as well as the U.S. Supreme Court's decisions in *Daubert* and *Kumho Tire*, do not provide the district court with the option to place so-called "limitations" on an expert witness's testimony or to edit that testimony into a form more palatable to the trial court judge. Judges have a binary choice under Rule 702: Admit the witness as an expert or not, recognizing that if you admit the witness as an expert, his or her expert opinion(s) come with them.

Some trial courts have imposed limitations on expert witnesses' testimony. These limitations typically prohibit an expert from testifying as to the credibility of another witness[546] or testifying outside the scope of their expertise.[547] Rule 704(b) also prohibits an expert witness from testifying as

[544] *See* FED. R. EVID. 702 (c)–(d).

[545] Gen. Elec. Co. v. Joiner, 522 U.S. 136, 146–47 (1997).

[546] *See* United States v. Adams, 271 F.3d 1236, 1245–46 (10th Cir. 2001).

[547] *See* United States v. Faines, 216 F. App'x 227, 229 (3d Cir. 2007). Here, an expert witness was called by the defense on the methodology of fingerprint comparison and examination. During testimony, the expert attempted to conduct a comparison of her own, but because she was not offered as an expert for that purpose, the trial court prohibited her from testifying to her opinion as to the comparison between the known and latent prints in the case. The Third Circuit affirmed the trial court's decision.

to his or her opinion about whether a defendant in a criminal trial "did or did not have a mental state or condition that constitutes an element of the crime charged or of a defense."[548] Such limitations are either explicitly or implicitly required under the Rules.[549] But the only tool that Rule 702 and *Daubert* provide to the courts is the authority to exclude expert witness testimony.[550] Unfortunately, the limitations imposed by the courts in recent firearms expert identification cases seek to do far more. They edit or make substantive and material changes to the expert's testimony, fundamentally altering or deleting the expert's opinion and substituting the judge's opinion instead. This goes well beyond the "gatekeeping" function envisioned for trial judges by the U.S. Supreme Court in *Daubert*. Here are some examples:

In *United States v. Tibbs*, the firearms identification expert opined that a .40 caliber cartridge case "was microscopically examined and identified as having been fired in [the recovered pistol], based on breechface marks and firing pin aperture shear marks."[551] This was an expert opinion of identification. The court altered this testimony by imposing "limitations" on the expert so that he could only testify that the recovered firearm cannot be "excluded" as having fired the recovered cartridge case.[552] This altered opinion was *not* the expert's opinion, only that of the court.[553] The trial court judge summarily rewrote the testimony of an expert witness, changing his

[548] FED. R. EVID. 704(b); *see also* United States v. Stahlman, 934 F.3d 1199, 1220 (11th Cir. 2019) (denying a defense expert from testifying as to the defendant's state of mind, where the expert would have opined the defendant did not intend to have sex with a minor and holding the expert testimony to be in violation of Rule 704(b)).

[549] Additionally, Rule 702 remains subject to Rule 403, which could be used by the courts to exclude otherwise admissible expert witness testimony if a judge found the probative value of said expert testimony was substantially outweighed by such factors as unfair prejudice, confusing the issues, or misleading the jury. However, the sole remedy under Rule 403 is not to limit an expert's testimony, but to exclude it.

[550] *See Joiner*, 522 U.S. at 146. In *Joiner*, the U.S. Supreme Court wrote: "We hold, therefore, that abuse of discretion is the proper standard by which to review a district court's decision to admit or exclude scientific evidence." *Id.* The abuse of discretion standard may be appropriate to review a trial court's decision to admit or exclude an expert's testimony. However, once that decision has been made by the district court, any "limitations" on said testimony are issues of law which should be reviewed *de novo* by the appellate courts.

[551] No. 2016 CF1 19431, 2019 D.C. Super. LEXIS 9, at *8–9 (D.C. Super. Ct. Sept. 5, 2019).

[552] *Id.* at *80–81.

[553] *Id.* One wonders how the court would have responded if the expert witness had refused to abide by the judge's limitations, believing that such alterations of his/her expert testimony would be untruthful or misleading.

"identification" opinion to one of "inconclusive."[554] In so doing, the jury was misled into believing that the firearms expert witness could not identify the cartridge case found at the crime scene as having been fired by the pistol recovered by law enforcement.[555] As discussed before, when a bullet or cartridge case can neither be identified nor excluded as having been fired by a particular firearm, that opinion becomes one of "inconclusive."[556] As applied, the judge's limitation in the *Tibbs* case had the practical effect of changing the expert's opinion testimony. This was not "gatekeeping" because the judge did not elect to exclude the expert from testifying. This was an abuse of discretion.

In *United States v. Shipp*, the district court judge had a firearms expert whose opinion was that both a bullet fragment and a cartridge case obtained at the crime scene were identified as having been fired by the recovered firearm, a Sig Sauer 9mm pistol.[557] However, the court imposed limitations that rewrote the examiner's testimony such that he could testify only that the marks on the test bullet and cartridge case fired by the recovered pistol were "consistent with" the bullet fragment and the cartridge case found at the crime

[554] *Id.* at *60; *see also* DOJ FIREARMS ULTR, *supra* note 28, at 3 ("[A]n 'inconclusive' conclusion is an examiner's opinion that there is an insufficient quality and/or quantity of individual characteristics to identify or exclude."). By prohibiting the examiner from testifying they could not identify nor exclude the cartridge case in the *Tibbs* case, the court has fundamentally altered the expert's opinion, changing it to one of "inconclusive."

[555] *Tibbs*, 2019 D.C. Super. LEXIS 9, at *77–81.

[556] *See* ASSOCIATION OF FIREARM AND TOOLMARK EXAMINERS, *AFTE Range of Conclusions*, 24 AFTE J. 233 (1992), https://afte.org/about-us/what-is-afte/afte-range-of-conclusions.

> Inconclusive:
>
> Some agreement of individual characteristics and all discernible class characteristics, but *insufficient for an identification*.
>
> Agreement of all discernible class characteristics without agreement or disagreement of individual characteristics due to an absence, insufficiency, or lack of reproducibility.
>
> Agreement of all discernible class characteristics and disagreement of individual characteristics, but *insufficient for an elimination*.

*Id.* (emphasis added). The Department of Justice defines the term "*inconclusive*" with regards to firearms examinations as follows: "'Inconclusive' is an examiner's conclusion that all observed class characteristics are in agreement but there is insufficient quality and/or quantity of corresponding individual characteristics such that *the examiner is unable to identify or exclude the two toolmarks as having originated from the same source*." DOJ FIREARMS ULTR, *supra* note 28, at 3 (emphasis added).

[596] United States v. Shipp, 422 F. Supp. 3d 762, 766–67 (E.D.N.Y. 2019).

scene.[558] The court directed the expert to testify that the 9mm pistol "[could not] be excluded" as the source for either the bullet fragment or the cartridge case but could not testify the 9mm pistol was the source of the bullet fragment or the cartridge case.[559] That was *not* the examiner's opinion. The judge instructed the expert witness on precisely what to say and how to say it. This was also not a so-called "limitation" imposed by the court, but a wholesale rewriting of an expert witness's testimony, again changing an "identification" opinion to one of "inconclusive." This, too, constitutes an abuse of discretion.

In *United States v. Davis*, the district court judge dealt with three firearms identification expert witnesses. These expert witnesses would render an opinion that (1) "certain bullets and casings found at one crime scene can be associated with bullets, spent shell casings, or firearms recovered from other crime scenes," and (2) that a caliber Smith & Wesson .40 caliber semi-automatic pistol was identified as the source of a cartridge case found at the murder scene.[560] The court "limited" the experts to testifying that marks made by the pistol were "consistent with" those on the recovered cartridge case.[561] None of the examiners were permitted to say the cartridge case was a "match" to the pistol, or to other cartridge cases from other crime scenes.[562] They were also not permitted to testify that all the cartridge cases they examined were fired by the same gun.[563] However, the opinion of these firearms examiners was not that the cartridge cases were "consistent with" anything; it was their opinion that they had *identified* the .40 caliber pistol as being the sole source of the fired cartridge cases.[564] The court's "limitations" erased their opinions and substituted its own beliefs about the weight of the expert's opinions in the case.

The judge in *United States v. Adams* went even further. Here, the firearms examiner identified a Taurus .40 caliber pistol recovered from the defendant's apartment as the source of several cartridge cases and one bullet found at the crime scene.[565] Initially, the court restricted the firearms examiner from opining about the identification of either the bullet or

[558] *Id.* at 783.

[559] *Id.* at 783–84.

[560] United States v. Davis, No. 4:18-cr-00011, 2019 U.S. Dist. LEXIS 155037, at *5–6 (W.D. Va. Sept. 11, 2019).

[561] *Id.* at 26.

[562] *Id.*

[563] *Id.* at 26–27.

[564] *Id.* at 5–6.

[565] United States v. Adams, 444 F. Supp. 3d. 1248, 1251, 1253 (D. Or. 2020).

cartridge cases, limiting the witness to testify that the Taurus pistol "could not be excluded" as the source of the ballistics evidence.[566] The court later revised its decision, prohibiting the expert from testifying about anything other than the class characteristics shared by the bullet and cartridge casings found at the crime and those made by test firings from the suspect's Taurus pistol. No testimony of individual characteristics was permitted, nor was the examiner allowed to provide his opinion about an identification or a match.[567] This was not the expert's true opinion. In fact, it is not an opinion at all. The witness was only allowed to testify about objective general rifling characteristics, which does not require a firearm examiner's expert testimony. If the court thought this witness was not reliable enough to offer an opinion as a firearms examiner, then why did the court permit him to be called as an expert?[568] Regardless, the "limitations" imposed by the court in *Adams* also constitute an abuse of discretion.

### *B. The Lack of Legal or Scientific Bases for Judicially Imposed Limitations*

A few skeptical judges have precluded firearms identification expert witnesses from using such terms as "identification" or "match."[569] Instead, these courts have ordered firearms examiners to dispense with their opinions and use terms such as "consistent with," "cannot be excluded," "could have fired," and "more likely than not."[570] These are *not* limitations, but substantive and material changes to the testimony of a witness.[571]

---

[566] *Id.* at 1256.

[567] *Id.* at 1267.

[568] *See id.* One of the reasons the judge in *Adams* may not have wanted to exclude the expert witness is because the district court judge realized that if he excluded the firearms expert, he would have run afoul of the ruling of the United States Court of Appeals for the Ninth Circuit in *United States v. Johnson*, where the court held the admission of firearms identification expert testimony was not an abuse of discretion. 875 F.3d 1265, 1281 (9th Cir. 2017). Given the Ninth Circuit's decision from just three years earlier, it seems unlikely a ruling by the district court to exclude such expert testimony in *Adams* would have been upheld by the Ninth Circuit on appeal.

[569] *See Adams*, 444 F. Supp. 3d at 1256; *Davis*, 2019 U.S. Dist. LEXIS 155037, at *26.

[570] United States v. Davis, No. 4:18-cr-00011, 2019 U.S. Dist. LEXIS 155037, at *24–27 (W.D. Va. Sept. 11, 2019).

[571] Indeed, it appears that attempts by judges to rewrite the opinion testimony of firearms expert witnesses may be futile anyway. In *Mock Jurors' Evaluation of Firearm Examiner Testimony*, Professors Brandon Garrett, Nicholas Scurich, and William E. Crozier conducted mock jury studies with 200 mock jurors, using the altered conclusions of firearms examiners authored by many of the courts recounted in this article. 44 L. & HUM. BEHAV. 412, 422 (2020). Their study found "that

An example of true limitations exists in DOJ's Firearms ULTR, where examiners still provide their opinion, unaltered by the court, yet subject to caveats such as not being able to declare they have identified the source of a fired bullet or cartridge case "to the exclusion of all other sources."[572] They cannot declare their opinion is with "absolute or 100% certainty."[573] Nor can they claim forensic firearms discipline has "a zero error rate."[574] These limitations are where the courts should focus, not on substantive or material changes to the expert's opinion. An example of genuine limitations which do not alter a witness's opinion are listed below, as are substantive or material changes to the expert's opinion, which the courts must strive to avoid and would constitute an abuse of discretion.

| **Examples of *limitations* of testimony** | **Examples of *substantive or material changes* to testimony** |
|---|---|
| No absolute or 100% certainty as to conclusions | "May have fired" |
| No identification opinions "to the exclusion of all other firearms" in the world. | "Consistent with" |
| No "zero error rate" for firearms discipline | "More likely than not" |
| No use of "a reasonable degree of scientific certainty" to weigh opinion | "Cannot be excluded" |
| No use of terms "unique" or "individualization" when referring to firearms | Class characteristics only; no discussion of individual characteristics |

many judicial and prosecution driven interventions to limit conclusion language for firearms testimony are not likely to be effective." *Id.*

[572] DOJ FIREARMS ULTR, *supra* note 28, at 3.

[573] *Id.* at 4.

[574] *Id.* at 3.

| | |
|---|---|
| Opinion not based on examining all other firearms in the world | Proposed testimony whose net effect alters the examiner's opinion from "identification" to "inconclusive" |

But where and how did these substantive and material changes originate? What are their bases in science or the law? Why can't the courts in these cases agree on a uniform standard of testimony or admissibility for firearms expert witnesses? To answer these questions, we must examine the etymology of the language used by these courts.

The genesis of court-ordered substantive and material changes to the opinion of firearms identification expert witnesses began with case of *United States v. Glynn*, where the district court judge directed the firearms expert witness not to use the word "match" and testify only that such a match of the ballistics evidence was "more likely than not."[575] This decision was unprecedented. Before the *Glynn* decision, no other state or federal court had ever used such language or terms to limit a firearms examiner. While the court in *Glynn* referred to the decisions of sister courts in *United States v. Monteiro*,[576] *United States v. Green*,[577] and *United States v. Diaz*,[578] none of these cases attempted to alter the substantive testimony of the firearms identification expert from testifying to the source of ballistics evidence or that a firearm was a "match" to a spent bullet or cartridge case. None of these cases ever discussed, let alone required, the term "more likely than not." Because no legal precedent existed for the use of this term, the *Glynn* decision constitutes a legal error committed by the Court. "And an error of law is necessarily an abuse of discretion."[579]

The term "more likely than not" also lacks any scientific or technical support. The 2009 NRC Report on "Strengthening Forensic Science" would not be released until a year after the *Glynn* decision, and the PCAST Report on forensic science would not exist for another eight years. The judge referenced the 2008 NRC Report on Ballistic Imaging in his decision;

[575] 578 F. Supp. 2d 567, 575 (S.D.N.Y. 2008).

[576] *See* 407 F. Supp. 2d 351, 375 (D. Mass. 2006).

[577] *See* 405 F. Supp. 2d 104, 120–22 (D. Mass. 2005).

[578] *See* No. CR 05-00167 WHA, 2007 U.S. Dist. LEXIS 13152, at *35–36 (N.D. Cal. Feb. 12, 2007).

[579] Adams v. Lab'y Corp. of Am., 760 F.3d 1322, 1331 (11th Cir. 2014).

however, that report made clear it was not to be used as a resource in determining the admissibility of firearms expert testimony, further reinforcing the court's error.[580] The only other reference the court made to any source was an affidavit submitted to the court by the aforementioned Professor Adina Schwartz,[581] a criminal law professor who possesses no scientific training and has been rejected as a potential expert witness in the area of firearms examination and identification by at least one federal district court.[582] The court did not cite a single published or peer-reviewed firearms study where firearms examiners evaluated or used the term "more likely than not." The term is not used or endorsed by AFTE, NIST, any forensic laboratory, or any forensic organization dealing with firearms examinations. The term "more likely than not" is bereft of any scientific or technical foundation. The *Glynn* court fabricated the term by means of an *ipse dixit*, fundamentally altering the opinion testimony of a firearms identification expert witness in the process.

The terminology contagion that originated in the *Glynn* decision has not spread far. Yet the magistrate judge in *United States v. Mouzone*, relying yet again on an affidavit and in-court testimony from Professor Adina Schwartz, would later find the words "more likely than not" were too generous.[583] The magistrate recommended to the district court judge that even this language was not to be used by the firearms expert to express any degree of certainty regarding his conclusions, a position later adopted by the district court

[580] BALLISTIC IMAGING, *supra* note 11, at 3. The report also states:

> We also note that the committee does not provide an overall assessment of firearms identification as a discipline nor does it advise on the admissibility of firearms-related toolmarks evidence in legal proceedings: these topics are not within its charge. The committee's charge is to determine the extent to which the toolmarks left on bullets and cartridge casings after firing a weapon can be captured by imaging technology. It is also to assess whether a ballistic image database—particularly a national RBID containing images of exhibits fired from all newly manufactured and imported guns—would be feasible and operationally useful, by which we mean capable of generating leads for follow-up and further investigation.

*Id.* at 3–4. " . . . [T]he proposal for this study explicitly precluded the committee from assessing the admissibility of forensic firearms evidence in court, either generally or in specific regard to testimony on ballistic imaging comparisons." *Id.* at 20.

[581] United States v. Glynn, 578 F. Supp. 2d 567, 569 (S.D.N.Y. 2008).

[582] *See generally* United States v. Taylor, 704 F. Supp. 2d 1192 (D.N.M. 2009).

[583] United States v. Willock, 696 F. Supp. 2d 536, 547 n.26, 574 (D. Md. 2010).

judge.[584] However, the firearms expert in *Willock* still testified to the identity or source of the cartridge cases at trial.[585] The magistrate judge in the *Mouzone*-*Willock* cases would go on to become a federal district court judge in another related case, *United States v. Medley*, eight years later.[586] The *Medley* case entailed another trial where the firearms examiner—who had testified more than eighty times as an expert witness in firearm identification[587]—opined that ballistics evidence from the crime scene was "identified" as coming from a firearm found in the defendant's possession.[588] In making his ruling, the judge relied extensively on the PCAST Report and his 2010 decision in *Willock*, orally entering his ruling into the record for two and a half hours[589] and at one point asking, "So, how do we try to square the circle?"[590]

The judge in the *Medley* case "squared the circle" in several ways. First, he declared the jury itself would have to make an examination of the ballistics evidence and authenticate it pursuant to Federal Rule of Evidence 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."[591] Then the court ruled the firearms examiner could testify that the marks on the cartridge case were "consistent with" the marks made by the defendant's gun on the test-fired cartridges.[592] However, he denied the firearms examiner the ability to opine that the two cartridge cases were fired by the *same* gun.[593] He prohibited the expert from using the word "identify," substituting his own term of "consistent with."[594] The prosecution informed the judge that such changes to the witness's testimony either contradicted or were not in keeping with either the AFTE Range of Conclusions or the laboratory's protocols, but the judge ignored

[584] *Id.*

[585] United States v. Mouzone, 687 F.3d 207, 216 (4th Cir. 2012).

[586] Transcript of Mot. Hr'g Proceedings at 1, United States v. Medley, 312 F. Supp. 3d 493 (D. Md. 2018) (No. 8:17-cr-00242-PWG).

[587] *Id.* at 78, 121.

[588] *Id.* at 112–13.

[589] *Id.* at 67–133. The transcript reveals the judge returned from a recess at 1:30 p.m. whereupon he began to discuss his ruling in the case until 3:57 p.m. *Id.*

[590] *Id.* at 86.

[591] *Id.* at 117–18.

[592] *Id.* at 119–20.

[593] *Id.* at 120.

[594] *Id.* at 126.

such objections, repeating what he expected the firearms expert witness to say and how he must say it.[595] The judge even provided the prosecution and expert witness a verbal script of how he believed the testimony should be conducted.[596]

The court's ruling in the *Medley* case also constitutes an abuse of discretion. While there is nothing wrong with a jury examining ballistics evidence on its own, the jury would not have the benefit of a high-quality, comparison microscope with an optical bridge for the review or the training and expertise to know how to interpret what they saw. The court realized such images would lack the precision of the images from a laboratory-grade comparison microscope; instead, the jury would have to use an HD-TV or printouts for comparison.[597] Nor would the jury be able to put the evidence into context with the aid of an expert witness to interpret it.

Once again, this is an instance of a court redacting the expert's opinion and substituting words of its own, changing the form and substance of an expert witness's opinion in contravention of laboratory protocols and published industry standards. The term "consistent with," which the court concocted, is not recognized by any scientific or technical body in the field of firearms examination. While "consistent with" appears in the AFTE Range of Conclusions, it is only in the context of either "identification" and/or "sufficient agreement."[598] AFTE does not use or recognize the term "consistent with" standing alone or isolated from an examiner's conclusions. By itself, the term has never been subjected to any peer-reviewed studies or validation studies. Neither the PCAST Report nor the 2009 NRC Report recommended usage of the language "consistent with" in lieu of an examiner testifying as to the source of a bullet or cartridge case. The court cites no legal, scientific, or technical basis for requiring the use of the term "consistent with."[599] The judge could not name a single forensic laboratory or a firearms expert that endorses or uses the term "consistent with" in reaching their conclusions.

---

[595] *Id.* at 124–25.

[596] *Id.* at 124–27.

[597] *Id.* at 128.

[598] *AFTE Range of Conclusions*, THE ASS'N OF FIREARM AND TOOLMARK EXAM'RS, https://afte.org/about-us/what-is-afte/afte-range-of-conclusions (last visited Mar. 26, 2020).

[599] Indeed, the only legal authority the judge could cite for use of the term "consistent with" was himself and his decision in the *Mouzone-Willock* cases from eight years before—y et another example of confirmation bias run amok in the judiciary. *See United States v. Willock*, 696 F. Supp. 2d 536, 574 (D. Md. 2009).

This substantive and material change matters. A jury is now left adrift to reason if the marks on the recovered cartridge cases are "consistent with" the firearm belonging to the defendant, and if so, how many other firearms it is "consistent with." The number is unknown, and all the court has done is sow the seeds of confusion. The court has committed an abuse of discretion because the judge was a party to the adulteration of the expert's testimony.[600]

Contrary to the judge's claims, this is not a so-called "limitation" of an expert witness's testimony. It is far more. This constitutes an arbitrary adulteration of an expert witness's opinion testimony. Under Federal Rule of Evidence 702, judges are charged with determining the admission of expert testimony and inquiring into its reliability.[601] As the U.S. Supreme Court said in *Daubert*, "The inquiry envisioned by Rule 702 is, we emphasize, a flexible one."[602] The inquiry may be flexible; however, the decision governing the admissibility of evidence is not. The court here possessed a binary option: Admit the expert testimony or exclude it. "Squaring the circle" by a trial court judge to substantially alter or make a material change to an expert's opinion testimony to conform to a judge's personal bias goes well beyond the "gatekeeping" function the U.S. Supreme Court had in mind when it decided *Daubert* or *Kumho Tire*.[603] The U.S. Supreme Court cautioned, "The focus, of course, must be solely on principles and methodology, not on the conclusions that they generate."[604] The *Medley* court focused on the conclusions of the firearms examiner and rewrote them to its own liking.[605] At best, this could be called an abuse of discretion.

Unfortunately, the *Medley* case was not the last time a court turned its back on Rule 702 in the area of firearms identification expert testimony. A year later, in *United States. v. Davis*, the U.S. district court judge cited the *Medley* decision and used the language "consistent with" to alter the

[600] *See* Hyundai Motor Am. v. Applewhite, 53 So. 3d 749, 758–59 (Miss. 2011) (holding that it was an abuse of discretion when the trial court judge failed to provide any relief to the opposing party when an expert witness attempted to make substantive and material changes to his deposition testimony in an errata sheet).

[601] FED. R. EVID. 702.

[602] 509 U.S. 579, 594 (1993).

[603] *See id.*; Kumho Tire Co. v. Carmichael, 526 U.S. 137, 149, 152, 158 (1999).

[604] *Daubert*, 509 U.S. 595.

[605] *See* Transcript of Mot. Hr'g Proceedings at 119–20, United States v. Medley, 312 F. Supp. 3d 493 (D. Md. 2018) (No. 8:17-cr-00242-PWG).

testimony of a firearms identification expert witness.[606] Like the *Medley* decision, the court cited no scientific, technical, or legal authority for the use of the term "consistent with."[607] In *United States v. Shipp*, another judge arbitrarily imposed the term "consistent with" on a firearms expert, citing the *Medley* decision as a legal foundation, though the court also permitted the firearms examiner to testify the firearm "cannot be excluded" as the source of the ballistics evidence.[608] Whatever their shortcomings, at least the rulings in *Davis* and *Shipp* followed another district court judge's ruling. In *United States v. Tibbs*, the district court judge created another form of terminology, ruling that the firearms expert could testify the firearm "may have fired" the cartridge case recovered from the crime scene and "cannot be excluded" as the source of the cartridge case.[609]

"More likely than not," "consistent with," "cannot be excluded," and "may have fired" are terms that judges manufactured with no reference to any science or technical bases. It is make-believe, not legal or scientific analysis. Neither Rule 702 nor the United States Supreme Court have given trial court judges the authority to rewrite an expert witness's opinion testimony, even if it is under the guise of "limitations." Even worse, these courts cannot seem to agree on a consistent or uniform language that is appropriate for firearms examiners. The D.C. Court of Appeals held in *Williams v. United States* that "it is plainly error to allow a firearms and toolmark examiner to unqualifiedly opine, based on pattern matching, that a specific bullet was fired by a specific gun."[610] But the court failed to give any guidance to either the lower courts or to firearms examiners as to what a "qualified" opinion in firearms and toolmarks examinations entails.[611] Regardless, it appears that the D.C. Court of Appeals is encouraging the lower courts to abandon Rule 702 and adopt undefined pre-admission criteria for determining the admissibility of certain expert witness testimony.

[606] No. 4:18-cr-00011, 2019 U.S. Dist. LEXIS 155037, at *26 (W.D. Va. Sep. 11, 2019) (where one of the firearms expert witnesses in the case, Scott McVeigh, was also the firearms examiner in *Medley*); Transcript of Mot. H'rg Proceedings, *supra* note 586, at 67–68.

[607] *See Davis*, 2019 U.S. Dist. LEXIS 155037, at *26.

[608] 422 F. Supp. 3d 762, 783 (E.D.N.Y. 2019).

[609] No. 2016 CF1 19431, 2019 D.C. Super. LEXIS 9, at *77–78 (D.C. Super. Ct. Sept. 5, 2019).

[610] 210 A.3d 734, 744 (D.C. 2019).

[611] *See id.*

### *C. Weight v. Admissibility*

The courts in *Adams*, *Tibbs*, *Shipp*, *Davis*, and *Williams* all admitted firearms-identification expert testimony. So, what was the purpose of these courts in going beyond the decision to admit such testimony and attempt to rewrite it instead? The answer can be found in *United States v. Glynn*, where the court wrote the following:

> [O]nce expert testimony is admitted into evidence, juries are required to evaluate the expert's testimony and decide what weight to accord it, but are necessarily handicapped in doing so by their own lack of expertise. There is therefore a special need in such circumstances for the Court, if it admits such testimony at all, to limit the degree of confidence which the expert is reasonably permitted to espouse . . . The problem is how to admit it into evidence without giving the jury the impression—always a risk where forensic evidence is concerned—that it has greater reliability than its imperfect methodology permits . . . [W]hen it comes to expert testimony, cross-examination is inherently handicapped by the jury's own lack of background knowledge, so that the Court must play a greater role, not only in excluding unreliable testimony, but also in alerting the jury to the limitations of what is presented.[612]

This was the reasoning enlisted by the court to rewrite the testimony of firearms expert witnesses: To impart the weight the firearms expert opinion testimony should have on the jury. It is also a demonstration of the trial court judge's abuse of discretion.

These courts have forgotten their limited role as "gatekeepers" under *Daubert* and Rule 702. The U.S. Court of Appeals for the Eleventh Circuit warned of such a lack of judicial discipline in *United States v. Barton*:

> Notwithstanding its critical gatekeeping function, the trial court is just that—a gatekeeper—and Rule 702 is a screening procedure, not an opportunity to substitute the trial court's judgment for that of a jury. In that regard, "it is not the role

[612] 578 F. Supp. 2d 567, 571–74 (S.D.N.Y. 2008).

> of the district court to make ultimate conclusions as to the persuasiveness of the proffered evidence."[613]

The Eleventh Circuit is far from alone in finding a lack of judicial discipline. For example, the U.S. Court of Appeals for the Federal Circuit ruled in *Apple, Inc. v. Motorola, Inc.*:

> A judge must be cautious not to overstep its gatekeeping role and weigh facts, evaluate the correctness of conclusions, impose its own preferred methodology, or judge credibility, including the credibility of one expert over another. These tasks are solely reserved for the fact finder. . . . That the gatekeeping role of the judge is limited to excluding testimony based on unreliable principles and methods.[614]

The U.S. Court of Appeals for the Seventh Circuit found similar issues with how trial court judges should perform their role as "gatekeepers" under Rule 702.[615] If, as the judge in the *Glynn* case claimed, the issues surrounding credibility were too difficult for a jury to decipher, the Seventh Circuit provided a solution for that: "If the judge believes expert testimony is too complex for the jury to appreciate important issues of reliability, such that admitting the testimony would prejudice the opposing party, the judge remains free to exclude such evidence under Rule 403."[616]

Note how the Seventh Circuit never mentioned placing so-called "limitations" on the purported expert witnesses' testimony in the event the reliability or credibility is too complex for the jury to appreciate. In addition, the U.S. Court of Appeals for the Sixth Circuit weighed in on this issue,

---

[613] 909 F.3d 1323, 1332 (11th Cir. 2018) (quoting Quiet Tech. DC-8, Inc. v. Hurel-Dubois UK Ltd., 326 F.3d 1333, 1341 (11th Cir. 2003)).

[614] 757 F.3d 1286, 1314–15 (Fed. Cir. 2014).

[615] Stollings v. Ryobi Techs., Inc., 725 F.3d 753, 765 (7th Cir. 2013) ("[S]oundness of the factual underpinnings of the expert's analysis and the correctness of the expert's conclusions based on that analysis are factual matters to be determined by the trier of fact.") (quoting Smith v. Ford Motor Co., 215 F.3d 713, 718 (7th Cir. 2000))); *id.* at 766 (noting that "an expert may . . . offer a conclusion that is subject to doubt").

[616] *Id.* at 766. The Seventh Circuit Court of Appeals is referring to Federal Rule of Evidence 403, "Excluding Relevant Evidence for Prejudice, Confusion, Waste of Time, or Other Reasons," which provides, "The court may exclude relevant evidence if its probative value is substantially outweighed by a danger of one or more of the following: unfair prejudice, confusing the issues, misleading the jury, undue delay, wasting time, or needlessly presenting cumulative evidence." FED. R. EVID. 403.

finding that any weakness in the factual bases of an expert's opinion bears on the weight of the expert's opinion, rather than its admissibility.[617]

In *Pyramid Techs., Inc. v. Hartford Cas. Ins. Co*., the U.S. Court of Appeals for the Ninth Circuit also dealt with a district court judge who abused his discretion in denying the admissibility of two expert witnesses.[618] In reversing the district court, the Ninth Circuit held: "After an expert establishes admissibility to the judge's satisfaction, challenges that go to the weight of the evidence are within the province of a fact finder, not a trial court judge. A district court should not make credibility determinations that are reserved for the jury."[619]

Finally, the U.S. Court of Appeals for the D.C. Circuit was perhaps the first appellate court to see a problem with district court judges overreaching in their "gatekeeper" role under Rule 702 and *Daubert*. It admonished judges to avoid weighing competing scientific studies and conflating questions of the admissibility of expert testimony, rather than the weight accorded such testimony by the jury.[620]

This is precisely what has happened with the handful of courts that have rewritten the testimony of firearms identification expert witnesses. Their attempts to subvert the role of the jury in weighing the certainty of expert witness testimony constitutes another abuse of discretion. Had those courts simply imposed limitations in line with those in the DOJ Firearms ULTR and left the matter at that, no one could object.

### *D. Getting It Wrong Most of the Time*

This article has laid out the range of error rates for the firearms and tool marks discipline, which ranges from less than 1% (calculated by your humble correspondent) to as high as 2.2% (calculated by PCAST). Reviewing the six studies cited by the PCAST Report and the three additional studies published since the PCAST Report, we see a reported false-positive error rate, ranging from 0% to approximately 1%. This suggests that when a firearms identification expert witness has "identified" or "matched" a bullet or

[617] *In re* Scrap Metal Antitrust Litig., 527 F.3d 517, 530–31 (6th Cir. 2008).

[618] 752 F.3d 807, 810–11 (9th Cir. 2014).

[619] *Id.* at 814. The Ninth Circuit also held, "Simply put, '[t]he district court is not tasked with deciding whether the expert is right or wrong, just whether his testimony has substance such that it would be helpful to a jury.'" *Id.* at 813 (quoting Alaska Rent-A-Car, Inc. v. Avis Budget Grp., Inc., 738 F.3d 960, 969–70 (9th Cir. 2013)).

[620] Ambrosini v. Labarraque, 101 F.3d 129, 141 (D.C. Cir. 1996).

cartridge case to a specific firearm, his or her opinion is highly likely to be correct about 99% of the time. Therefore, it is axiomatic that when a court rewrites a forensic examiner's testimony which is *right* nearly 99% of the time, the trial court judge is getting it *wrong* nearly 99% of the time.

The scenario has played out on several occasions in cases such as *Glynn* and its progeny in the *Adams*, *Medley*, *Tibbs*, *Shipp*, *Davis*, and *Williams* cases. This unauthorized practice constitutes an abuse of discretion.

The United States Court of Appeals for the Eleventh Circuit has described the trial court judge's proper role:

> Once an expert opinion has satisfied *Daubert*, a court may not exclude the opinion simply because it believes that the opinion is not—in its view—particularly strong or persuasive. The weight to be given to admissible expert testimony is a matter for the jury . . . It is not the role of the district court to make ultimate conclusions as to the persuasiveness of the proffered evidence.[621]

The United States Supreme Court foresaw a day when a trial court judge might question the opinion testimony of an expert witness, and wonder whether a jury might be incapable of determining the credibility of such opinion testimony. The Court held that:

> In this regard respondent seems to us to be overly pessimistic about the capabilities of the jury and of the adversary system generally. Vigorous cross-examination, presentation of contrary evidence, and careful instruction on the burden of proof are the traditional and appropriate means of attacking shaky but admissible evidence.[622]

Let us hope the trial courts will follow the wisdom espoused by the United States Supreme Court over a quarter century ago.

## XIII. CONCLUSION

On November 15, 2014, police responded to a 911 call and found the bodies of three people shot to death in Room 149 at the Economy Inn in

[621] Seamon v. Remington Arms Co., 813 F.3d 983, 990 (11th Cir. 2016).

[622] Daubert v. Merrell Dow Pharms., Inc., 509 U.S. 579, 596 (1993).

Springfield, Missouri.[623] A fourth victim died of his wounds five days later.[624] On November 30, Scott Goodwin-Bey entered a convenience store in Springfield carrying a Ruger 9mm semi-automatic pistol.[625] The store clerk took the weapon from Goodwin-Bey and called police, who seized the pistol and arrested Goodwin-Bey.[626]

Ballistics examinations revealed the thirteen cartridge cases and eleven bullets recovered from both the crime scene and the victim's bodies were identified by firearms examiners as originating from the same Ruger 9mm pistol carried by Goodwin-Bey.[627] Additionally, an eyewitness claimed to have seen the shooting and identified Goodwin-Bey as the shooter.[628]

The Missouri state trial court held a pre-admissibility hearing regarding the firearm expert's testimony in court.[629] It marked the first time a judge considered the PCAST Report in weighing the admissibility of firearms identification expert witness testimony.[630] The court issued its decision on December 16, 2016, where it compared firearms identification evidence to the results of a polygraph examination, then declared "this Court very reluctantly will allow the State's lab person to testify, but only to the point this gun could not be eliminated as the source of the bullet."[631]

Unable to pursue the case for want of admissible expert testimony which could identify Goodwin-Bey's pistol as the murder weapon, the prosecution

[623] Jess Rollins, *Inside Room 149: Man told SGF police he saw quadruple homicide*, SPRINGFIELD NEWS-LEADER (Feb. 9, 2015, 5:39 PM), https://www.news-leader.com/story/news/crime/2015/02/09/man-charged-connection-quadruple-homicide-springfield-hotel/23122707/.

[624] *Id.*

[625] *Id.*

[626] *Id.*

[627] *Id.*

[628] *Id.*

[629] United States v. Goodwin-Bey, 718 F. App'x 447, 447–48 (8th Cir. 2018) (allowing state expert to testify). This case is no longer good law, having been superseded by the Missouri Court of Appeals in two subsequent decisions. *See* State v. Boss, 577 S.W.3d 509, 519 (Mo. Ct. App. 2019); State v. Mills, 623 S.W.3d 717, 732 (Mo. Ct. App. 2021).

[630] Jordan Smith, *Advocates Challenge Mysterious Justice Department Statement That Undercuts Forensic Science Reform*, THE INTERCEPT (Aug. 8, 2021, 6:00 AM), https://theintercept.com/2021/08/08/forensic-science-reform-justice-department/.

[631] State v. Goodwin-Bey, No. 1531-CR00555-01, slip op. at 6–7 (Mo. Cir. Ct. Dec. 16, 2016).

dismissed all four murder charges against Goodwin-Bey.[632] To date, the killer of Trevor Fantroy, Lewis Green, Danielle Keyes, and Christopher Freeman has not been brought to trial.[633]

The FBI recently released its Uniform Crime Report for 2020, showing the number of homicides in the United States jumped from 16,669 in 2019 to 21,570 in 2020.[634] This was an increase of almost 30% and the largest annual increase in homicides since record-keeping began in the 1960s.[635] Approximately 72% of all homicides in the U.S. are committed with some kind of firearm.[636] Given this fundamental threat to public health and safety, it seems absurd that some courts would attempt to diminish or exclude firearms identification expert testimony, which is often central to the prosecution of such cases. This is especially true when one considers the firearms identification discipline has less than a 1% false-positive rate and a demonstrated history of reliability.

The PCAST Report shed no light on the firearms and toolmark discipline; rather, it needlessly cast a shadow on a proven forensic technique, leaving judges and prosecutors in the dark about the true reliability of firearms identification expert testimony. Repeated empirical studies continue to

[632] KMBC News, *Charges dropped against Missouri quadruple homicide suspect*, THE ASSOCIATED PRESS (Dec. 29, 2016, 7:55 AM), https://www.kmbc.com/article/charges-dropped-against-missouri-quadruple-homicide-suspect/8545055.

[633] Goodwin-Bey was convicted of being a felon in possession of a firearm in violation of 18 U.S.C. §§ 922(g)(1) and 924(a)(2) in the United States District Court for the Western District of Missouri for his possession of the 9mm Ruger pistol at the convenience store. *Goodwin-Bey*, 718 F. App'x at 447–48. When arrested, he was driving the same white Lincoln Town Car described by the eyewitness to the murders at the Economy Inn two weeks before. *Id.* at 448; Jess Rollins, *Inside Room 149: Man told SGF police he saw quadruple homicide*, SPRINGFIELD NEWS-LEADER (Feb. 9, 2015, 5:39 PM), https://www.news-leader.com/story/news/crime/2015/02/09/man-charged-connection-quadruple-homicide-springfield-hotel/23122707/. Goodwin-Bey was sentenced to prison with a projected release date of June 7, 2023. *See Inmate Locator*, FED. BUREAU OF PRISONS, https://www.bop.gov/inmateloc/ (last visited Oct. 30, 2021).

[634] Ryan Lucas, *FBI Data Shows an Unprecedented Spike in Murders Nationwide in 2020*, NAT'L PUB. RADIO (Sept. 27, 2021, 1:12 PM), https://www.npr.org/2021/09/27/1040904770/fbi-data-murder-increase-2020.

[635] *Id.*; *see also* Josiah Bates, *FBI Data Shows a Surge in Murders in 2020. That's Not the Full Story*, TIME MAG. (Sept. 30, 2021, 12:32 PM), https://time.com/6102149/fbi-homicide-stats-analysis/.

[636] The 2017 FBI crime statistics report cites 15,206 homicides, with 11,014 committed by a firearm. *Murder Victims by Weapon*, FBI UNIF. CRIME REP., https://ucr.fbi.gov/crime-in-the-u.s/2019/crime-in-the-u.s.-2019/tables/expanded-homicide-data-table-8.xls (last visited Dec. 27, 2021). This amounts to show that approximately 72% of all homicides are committed by firearms.

demonstrate that the firearms identification discipline is remarkably accurate in identifying a particular bullet or cartridge case as having been fired by a specific firearm. This review of the 2016 PCAST Report and the case law that followed culminates in two recommendations for courts and practitioners.

First, courts should recognize the long-standing reliability of the firearms identification discipline and the examiners who testify to that discipline. As firearms examiners are correct nearly 99% of the time, their opinion testimony should be admitted without substantive or material changes to the examiner's opinion. Courts should abandon attempts to rewrite firearm examiner's opinions to conform to their subjective lay beliefs, as doing so could potentially mislead the jury. Judges would be wise to understand the lack of legal or scientific foundation supporting the PCAST Report and reject the flawed scientific pronouncements it made, just as Attorney General Loretta Lynch has done. "Limitations" to any expert's testimony should never substantively or materially alter an expert's opinion. However, if a particular examiner is found unreliable by a judge, then exclusion of the expert witness testimony may be the court's only recourse.[637]

Second, prosecutors should oppose, object to, and appeal any attempt by trial court judges to rewrite, alter, amend, or exclude the opinions and testimony of firearms identification expert witnesses. Because these alterations make substantive and material changes to a firearm expert's opinion, they constitute an abuse of discretion. Furthermore, because court-ordered "limitations" concern questions of law and not the admissibility of evidence, appellate courts should review *de novo* all attempts at rewriting or "limiting" an expert's opinion testimony.

It is a lamentable day for science and the law when people in black robes attempt to substitute their opinions for those who wear white lab coats.

[637] Recall that in the *Baldwin (Ames) Study*, "[a]ll but two of the 22 false identification calls were made by five of the 218 examiners, strongly suggesting that this error probability is not consistent across examiners." *Baldwin Study*, *supra* note 211, at 16. Had the results of the five examiners who made those errors been removed from the findings, the false positive rate for the remaining 213 firearms examiners who participated in the study would have been 0.00091 or 0.09%.