**SUPERIOR COURT OF THE DISTRICT OF COLUMBIA**
**Criminal Division – Felony Branch**

| | | |
|---|---|---|
| **UNITED STATES** | : | **Case No.: 2021 CF1 000968** |
| | : | |
| **v.** | : | **Hon. Marisa Demeo** |
| | : | |
| **DELONTA STEVENSON** | : | |

**GOVERNMENT'S OPPOSITION TO DEFENDANT'S MOTION TO PRECLUDE, OR IN THE ALTERNATIVE LIMIT FIREARM AND TOOLMARK IDENTIFICATION TESTIMONY**

Defendant, Delonta Stevenson ("defendant") filed a motion to exclude or limit expert firearm and toolmark testimony pursuant to Rule 702 and *Daubert v. Merrell Dow Parm. Inc.*, 509 U.S. 579 (1993). Def's Mtn. at 1. The Court should deny the defendant's Motion. A review of the scientific data presented in this pleading, which primarily focuses on post-PCAST[1] studies, establishes that firearms and toolmark examiners reliably make source conclusions involving firearms, cartridge casings, and bullets with low rates of error. The defendant has failed to provide any information to undermine this conclusion. Rather, the defendant primarily relies on the findings of three outdated policy reports: 1) the 2008 Ballistic Imaging Report (2008 NAS Report); 2) the 2009 Report by the National Academy of Sciences (2009 NAS Report); and 3) the 2016 PCAST Report. Similarly, the cases cited by the defendant, including *United States v. Tibbs*, 2016 CF1 019432, 2019 WL 4359486 (D.C. Super. Ct. September 5, 2019), either did not consider or did not have the benefit of the substantial body of post-PCAST research and thus are outdated and incomplete. As such, the Court should deny the defendant's Motion, and admit the expert conclusions as described herein.

[1] "PCAST" or "2016 PCAST Report" refers to Report to the President, Forensic Science in Criminal Courts: Ensuring Scientific Validity of Feature-Comparison Methods, September 2016. *available at* https://obamawhitehouse.archives.gov/sites/default/files/microsites/ostp/PCAST/pcast_forensic_science_report_final.pdf (accessed 5/7/2023).

## BACKGROUND

### A Firearms and Toolmarks Identification[2]

Firearm identification has been a forensic discipline since the 1920s. *See* The History of Firearm and Toolmark Identification, AFTE Journal, 1999 Volume 31, Number 3 (Summer), pp. 266-284. One of the first laboratories was associated with Northwestern University in the 1930s. *Id.* Firearms identification is a subset of the broader forensic discipline known as toolmark identification. Toolmark examiners are trained to examine the marks left by tools on a variety of surfaces in an attempt to associate a toolmark to a particular tool that made the mark. F irearms, which are designed to launch projectiles by means of combustion, are comprised of component tools that impart marks on bullets and cartridge cases.[3] *See* AFTE.org/resources/swggun-ark, Foundational Overview of Firearm/Toolmark Identification.

A firearm imparts different types of marks on the various components of a cartridge. With respect to bullets, cuts within a gun barrel ("grooves") and raised surfaces ("lands") create corresponding depressed "land impressions" and raised "groove impressions" as bullets travel through a barrel. The twist imparted on a bullet may be either left or right, depending on the direction of the lands and grooves. With respect to cartridge casings, contact between the cartridge and the breech create "breech face marks," and the impact of the firing pin on the primer creates a "firing pin impression" on the primer itself. Visual representations of these areas appear below:

---

[2] This section was written with the assistance of and reviewed by Todd Weller, an expert in firearms and toolmark examination with more than 20 years of forensic science experience. A summary of Mr. Weller's qualifications appears in Govt. Ex. 2. A negative finding related to Mr. Weller from a court in Cook County, Illinois is described in footnote 16. Additionally, this section is based in part on the work of Stephen Bunch (formerly the Unit Chief with the FBI Firearms/Toolmark Unit), FBI Firearms Analyst Douglas Murphy, ATF firearms examiner Greg Klees, and John Murdock.

[3] A bullet is the projectile that is fired from a rifled firearm. A cartridge case/casing or casing is the container for all the other components that comprise an unfired round of ammunition (i.e., cartridge). Together a bullet and cartridge casing make up an unfired round of ammunition or cartridge.



The working edges of tools, which include components of firearms that contact ammunition, generally consist of some type of hard material, such as steel, to ensure strength and durability of the tool while work pieces are generally made of softer materials (ammunition metals

typically consist of copper, lead, and brass, which are softer than steel). These tool surfaces contain random, microscopic irregularities produced during the tool's manufacture and/or through subsequent wear through use and abuse. These irregularities, which are formed randomly, can individualize or distinguish one tool from another. Because these irregularities or individual characteristics[4] are typically imparted by contact onto the work piece, the comparative study of the imparted markings allow the tool to be individually associated or identified as having produced the mark. The presence, observation, and comparison of these random toolmarks on tools form the hypothetical propositions upon which the discipline of toolmark identification is based.

Firearm and toolmark identification is based upon two testable propositions:

> *Proposition #1:*
> Toolmarks imparted to objects by different tools will rarely if ever display agreement sufficient to lead a qualified examiner to conclude the objects were marked by the same tool. That is, a qualified examiner will rarely if ever commit a false positive error (misidentification).
>
> *Proposition #2:*
> Most manufacturing processes involve the transfer of rapidly changing or random marks onto work pieces such as barrel bores, breech faces, firing pins, screwdriver blades, and the working surfaces of other common tools. This is caused principally by the phenomena of tool wear and chip formation or by electrical/chemical erosion. Microscopic marks on tools may then continue to change from further wear, corrosion, or abuse.

*See* Bunch S., Smith E, Grioux B., Murphy D., *Is a Match Really a Match? A Primer on the Procedures and Validity of Firearm and Toolmark Identification*, FORENSIC SCI. COMMUNICATIONS, July 2009, Vol 11, No. 3 (*available at https://archives.fbi.gov/archives/about-us/lab/forensic-science-communications/fsc/july2009/review/2009_07_review01.htm* ("Bunch et. al.").

Examiners are trained to recognize and evaluate the following characteristics: (1) class

[4] Throughout this Opposition "individual characteristics" and "random characteristics" are used interchangeably. Mr. Monturo will use the term "random characteristics" during his testimony.

characteristics; (2) subclass characteristics; and (3) individual characteristics. Class characteristics result from design features and are predetermined during the manufacturing process. *See* ASSOCIATION OF FIREARM & TOOL MARK EXAMINERS, AFTE GLOSSARY 38 (6th ed. 2013) ("AFTE Glossary").[5] These are measurable features that indicate a restricted group source.[6] For a fired bullet, class characteristics include caliber and the number, direction, and width of the land and groove impressions. For a cartridge case, class characteristics typically include (1) caliber; (2) the firing pin impression on the primer, which can appear in various shapes (including circular, rectangular, hemispherical, and elliptical); and (3) the shape of the firing pin aperture and the type of breach face impression, which can be in different shapes and orientations (e.g., arched, circular, parallel, etc.). Even one type of class characteristic difference between a bullet or cartridge casing from a crime scene and a firearm (or between bullets and cartridge cases from a crime scene) will result in an elimination.

Subclass characteristics are more restrictive than class characteristics and are consistent among items manufactured by the same tool in the same approximate state of wear. These characteristics can exist within a particular production run in the manufacturing process and occasionally arise from (1) imperfections in a machine tool that persist during the production of multiple firearm components; or (2) extreme hardness differences between the machine tool and the work pieces. Unlike class characteristics, they are not common to all units of a particular make

[5] The AFTE Training Manual, Technical Procedures Manual and Glossary are too large to be attached to this pleading. They are publicly available for download at www.AFTE.org.

[6] Although class characteristic consistency narrows the pool of alternative firearms down somewhat, it still leaves open potentially 100s or 1000s of firearms that could have left the class marks on the evidence. "Although the agreement in class characteristics indicates a restricted group source and often is highly probative information, currently it is not possible to determine quantitatively the probative value of a level-one, class-only match. Firearm production and ownership figures are constantly changing and difficult to obtain." *See* Bunch et. al. As discussed herein, individual/random characteristics narrow that pool to substantially far fewer (to a few or one).

and model of firearm. Unlike individual characteristics, they continue throughout a period of manufacturing. The profession has been aware of the importance of subclass characteristics and documented their existence as early as the 1930s. The discipline's knowledge regarding the existence and features of subclass characteristics is thoroughly documented in the AFTE Journal. *See* Nichols, R., *Subclass Characteristics: From Origin to Evaluation*, AFTE JOURNAL 50(2) (2018) pp. 68-88 ("Nichols Subclass Review") (Exhibit 4c). Nichols Subclass Review discusses the definition, origin, and evaluation of subclass characteristics, citing to over 100 articles many of which were authored by toolmark examiners and published in the AFTE Journal. The profession has defined, studied, and published information for examiners regarding the proper identification and use of subclass characteristics and qualified examiners are trained to distinguish subclass characteristics from individual characteristics because identifications may not be made from subclass characteristics.

Individual or random characteristics, on the other hand, consist of microscopic, random imperfections in the barrel or firing mechanism created by the manufacturing process, wear, corrosion, or abuse. Individual characteristics are unintended microscopic features that occur due to the machining process and random chip formation during manufacturing. Individual characteristics typically fall into two categories: (1) striated marks made by movement of the tool parallel to the work piece (for example a bullet (workpiece) traveling through a barrel (the tool)), and (2) impressed marks made by the tool pressing (perpendicularly) into the work piece. A fired bullet usually has striated marks. A cartridge casing, on the other hand, can have both impressed and striated marks. Once the firearm is fired, striated marks may also be imparted to the cartridge case wall (side), and impressed marks may be imparted to the cartridge case by the gun's firing pin and breech. With semi-automatic weapons, additional marks can be made as the cartridge case is

expelled from the gun. In general, the tool working surfaces in a firearm can slowly change over time from wear. "For bullets fired through a barrel in sequential fashion, bullet number 1 usually will display significant microscopic correspondence to bullet number 10, 100, or even number 500 or 1000 but, depending on the firearm and caliber, may not achieve this for bullet number 50,000. Impressed marks are more persistent by their nature. Given relatively clean parts, firing pin and breechface impressions at levels of microscopic significance can persist for many thousands of firings." *See* Bunch et al (citing Bonifant and Dekinder 1999). The absence of infinite persistence does not diminish the reliability of examiner conclusions or the field of firearms and toolmarks. *Id.* As microscopic similarities diminish, the likelihood of an inconclusive result increases, but the likelihood of a false positive will remain unchanged, *id.*, or drop, *see* Stanley J. Bajic, *Report: Validation Study of the Accuracy, Repeatability, and Reproducibility of Firearm Comparisons*, AMES LABORATORY-USDOE TECHNICAL REPORT # ISTR-5220, p.55, Table XXI, (2020) ("Ames II") (Govt. Ex. 7).

Since the inception of firearms and toolmark identification as a forensic discipline, firearms examiners have been using a method known as "pattern matching" to determine whether sufficient similarity exists between toolmarks to warrant a conclusion that two bullets or two cartridge cases came from the same firearm. In 1992, AFTE memorialized the Theory of Identification in an attempt to summarize the basis of opinions of common origin in toolmark comparisons:

> 1. The theory of identification as it pertains to the comparison of toolmarks enables opinions of common origin to be made when the unique surface contours of two toolmarks are in "sufficient agreement."
>
> 2. This "sufficient agreement" is related to the significant duplication of random toolmarks as evidenced by a pattern or combination of patterns of surface contours. Significance is determined by the comparative examination of two or more sets of surface contour patterns comprised of individual peaks, ridges and furrows. Specifically, the relative height or depth, width, curvature and spatial relationship of the individual peaks,

> ridges and furrows within one set of surface contours are defined and compared to the corresponding features in the second set of surface contours. **Agreement is significant when agreement in individual characteristics exceeds the best agreement demonstrated between toolmarks known to have been produced by different tools and is consistent with agreement demonstrated by toolmarks known to have been produced by the same tool.** The statement that "sufficient agreement" exists between two toolmarks means that 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.
>
> 3. Currently the interpretation of individualized/identification is subjective in nature, founded on scientific principles and based on the examiner's training and experience.

*See* AFTE Criteria for Identification Committee, *Theory of Identification, Range of Striae Comparison Reports and Modified Glossary Terms – An AFTE Criteria for Identification Committee Report*, 24 AFTE JOURNAL 336, 336–40 (1992) (emphasis supplied); AFTE GLOSSARY at 138.

The bolded portion of the AFTE Theory of Identification encapsulates how an examiner learns, through training and ongoing experience, the amount of microscopic (tool)mark agreement expected in same source (i.e. known match) versus difference source (i.e. known non-match) samples. The AFTE Theory of Identification specifically directs the profession to seek out and use worst case scenario ("best known non-match") samples as part of training. Examiners learn to recognize the amount of agreement associated with worst case scenarios, or "best known non-matches," by examining fired bullets and cartridge cases collected from samples such as sequentially manufactured firearms where the similarity in (tool)marks from one firearm (or tool) to the next is at its greatest because the firearms are taken off the manufacture line one after another.

The discipline of firearm and toolmark identification examination methodology can be summarized as involving two phases of analysis. The first phase focuses on the classification and evaluation of existing class characteristics. If these are in agreement, the examination moves to

the second phase involving the comparative microscopic evaluation of individual characteristics. This culminates with an opinion decision concerning sufficient agreement or source[7]. In addition, firearm and toolmark examination includes quality assurance measures to ensure reliable results, including the documentation of examination and verification of results.

The following outlines the methodology firearm and toolmark examiners follow to determine a common source ("ECCV"):

***Evaluation[8]:*** The initial examination phase evaluates evidence to determine if the observed class characteristics are the same between two specimens (two unknown specimens, or an unknown and known specimen). If the specimens are suitable for examination and the class characteristics are consistent, then it is possible that the toolmarks were produced by the same tool (such as a firearm). If at least one class characteristic is different, then the two specimens can be eliminated as having been produced by the same tool.

---

[7] This decision involves some degree of subjectivity. However, it is bound by concepts discussed herein upon which all examiners are trained, such as the best agreement demonstrated between toolmarks known to have been produced by different tools and consistent with agreement demonstrated by toolmarks known to have been produced by the same tool. Additionally, one white box study, discussed *infra* at.47-49, shows that toolmark examiners, even when from different laboratories (15), each working independently, are mostly using the same amount and same location of microscopic marks when concluding identification. *See* 1/12/22 Weller Decl. ¶ C27; Pierre Duez et al., *Development and Validation of a Virtual Examination Tool for Firearm Forensics*, 63 J. FORENSIC SCI. 1069, 1069-1084 (2018). A second white box study, in which 40 different tests sets were independently compared by approximately 30 different examiners, concluded, in part: "Examiners had a high amount of agreement with regard to the areas useful for identification and elimination as well as those areas which should be avoided for definitive source attribution. It is worth reiterating that examiners worked independently and that the described annotation map patterns emerged when these independent submissions were combined. This consistency reinforces the fact that examiners typically agree on the toolmarks most important and most reliable for reaching source conclusions." *See* 1/12/22 Weller Decl. ¶ C28; Chad Chapnick, et al. "Results of the 3D Virtual Comparison Microscopy Error Rate (VCMER) Study for firearm forensics" 66 J. FORENSIC SCI. 569, 557-570 (2020) DOI: 10.1111/1556-4029.14602 p 1-14.

[8] This step may be broken into two steps, the first being evaluation of the ballistics evidence submitted as part of the case. If some or all of the evidence is suitable for further examination, an examiner will move on to classify the class characteristics of each item and any agreement among each item that is part of a case (i.e. bullets/cartridges). If broken into two steps, the acronym would be ECCCV.

***Comparison:*** If the class characteristics are consistent between two specimens, then a comparative examination is performed using a comparison microscope. This comparison is conducted to determine: 1) if any marks present are subclass characteristics and/or individual characteristics, and 2) the amount of correspondence of any observable individual/random characteristics.

***Conclusion:*** If sufficient agreement of individual/random characteristics is observed between two specimens, an identification conclusion is rendered. If all of the discernible class characteristics are the same, but sufficient agreement of the individual characteristics is not observed, an inconclusive result is rendered. In some situations, an elimination conclusion may be rendered based on observed differences in individual/random characteristics.

***Verification (Quality Assurance):*** A verification process is employed to ensure proper conclusions are rendered. As outlined in a laboratory's quality assurance policy, a mechanism should be in place to determine which cases will require verification. Many laboratories require verification of all identifications. *See* afte.org/resources/swggun-ark/summary-of-the-examination-method.

Using this methodology for examining tool-marked surfaces, there are four conclusions that examiners reach when conducting an examination: (1) identification, (2) inconclusive, (3) elimination, and (4) unsuitable for comparison. Examiners undergo standardized technical training designed to develop cognitive skills to recognize patterns of individual characteristics necessary to make an identification. However, there is no way to be *absolutely* (100%) certain of any identification without comparing a particular set of marks to marks created by every firearm produced since the invention of the modern-day firearm. This would be an impossible endeavor. *See* Todd Weller January 12, 2022 Decl. ¶ H4 (Govt. Ex. 2) ("1/12/22 Weller Decl."). Thus, an

examiner cannot rule out with absolute certainty the highly unlikely event that two different firearms will produce indistinguishable individual characteristics.

As discussed below, the field of forensic firearm and toolmark identification continues to undergo testing in the form of (1) technical research; (2) validation studies; and (3) proficiency testing. Validation studies are the most comprehensive way to test and validate firearms and toolmark identification as a reliable forensic science. These tests involve "ground truth," so it is known with absolute certainty from where each of the test components came. Using the same methods and identification criteria as those in actual casework, qualified examiners have consistently reached correct conclusions with exceptionally low error rates. *See* 1/12/22 Weller Decl. ¶¶ C1-C38. Even where researchers have studied bullets and cartridge cases fired from consecutively manufactured firearms – where the possibility of a false-positive conclusion is at its highest – trained examiners have been able to readily distinguish marks produced by the various firearms.

**B Summary of Facts**

The evidence in the case will established the following series of events. On January 18, 2021, just prior to the shooting at issue, which occurred at approximately 11:18 a.m., Troy Williams asked the decedent, Terrance Allen, for a ride to Stanton Glenn Apartments where he planned to check on his mother. Mr. Allen obliged. Mr. Allen's friend, James Fye, accompanied them so he and Mr. Allen could run errands later in the day. As Mr. Williams, Mr. Allen, and Mr. Fye drove in a white Ford Crown Victoria towards the entry/exit gate of the Stanton Glenn Apartments, Defendants Thomas and Stevenson sped towards them in a green Volvo SUV. Prior to getting in the vehicle, Defendant Thomas was captured on surveillance video removing a rifle from the green Volvo and Defendant Stevenson was captured on surveillance video holding a dark

colored handgun. Defendant Thomas was driving. Defendant Stevenson was seated in the front passenger's seat. Once in the car, Defendant Stevenson was armed with the rifle. As Defendants Thomas and Stevenson drove towards the car occupied by Mr. Williams, Mr. Allen, and Mr. Fye, Defendant Stevenson opened fired at the three men with the rifle.

Under fire, Mr. Williams fired back once in defense. Additionally, an on-duty SPO stationed at the Stanton Glenn entry/exit gate at the time of the shooting saw what was occurring and fired his firearm at the green Volvo SUV. Officers recovered the firearm fired by Mr. Williams – a Taurus G2 9mm firearm (CS1-43)–from Mr. Williams. Officers also recovered the SPO's firearm – a Smith & Wesson model M&P caliber 9mm semiautomatic pistol (P1-4).

After firing at least thirty-four rifle rounds at the white Ford Crown Victoria occupied by the three victims, Defendants Thomas and Stevenson sped away. Hearing the shots and seeing the green Volvo SUV fleeing the scene at a high rate of speed, an unmarked MPD unit pursued the green Volvo SUV. The unmarked MPD unit pursued the SUV to the 2600 block of 29th Street, Southeast where the green Volvo SUV crashed. Defendants Thomas and Stevenson got out of the crashed Volvo and ran making good their escape. Witnesses observed two individuals fleeing the vehicle that crashed and heading northbound up 29th Street towards a wooded path. Officers later recovered an Omni Hybrid semiautomatic rifle (CS3-1) in Defendants Thomas and Stevenson's flight path.

At approximately 11:10 a.m., MPD officers responded to the Stanton Glenn Apartments, 3040 Stanton Road Southeast, to address the gunshots that resulted in the murder of Mr. Allen. Upon arrival, they found Terrance Allen inside of a car. He was unconscious, unresponsive, and suffering from a gunshot wound to the head. He was rushed to George Washington University Hospital where he died. Officers found James Fye and Troy Williams suffering from non-life-

threatening gunshot wounds. They were taken to separate hospitals for treatment and released.

Law enforcement recovered the following firearms related evidence from the shooting/homicide crime scene:

- Three (3) SPEER 9mm Luger cartridge casings (CS1-1 through CS1-3);
- Thirty-four (34) PSD 17 .223/5.56 cartridge casings (CS1-4 through CS1-8 and CS1-11 through CS1-39); and
- One (1) bullet fragment (CS1-40).

Law enforcement recovered the following firearms related evidence from the white Ford Crown Victoria the decedent and surviving victims occupied at the time of the shooting:

- Six (6) sets of projectile fragments (MV1-1; MV1-3; MV1-7; MV1-8; MV1-10; and MV1-11);
- One (1) RP 9mm Luger casing (MV1-4) from the rear driver side floor under a mat;
- One (1) RP 9mm Luger cartridge (MV1-5) from the rear passenger side seat; and
- One (1) projectile and projectile tips (MV1-9).

Law enforcement recovered the following firearms related evidence from Defendant Thomas' green Volvo SUV:

- One (1) bullet (MV3-9) from the front passenger floor.

Law enforcement recovered the following firearms related evidence from the autopsy of Terrance Allen:

- Four (4) projectiles/metal bullet cores (A1-4; A1-5; A1-7; and A1-8) recovered from the decedent's body;
- One (1) projectile fragment/metal fragment (A1-3) recovered from the decedent's body; and
- Three (3) projectile fragments/metal fragments (A1-6) recovered from the decedent's body.

A NIBIN lead linked the recovered rifle (CS3-1) to a Destruction of Property on August 23, 2020, involving Andrea Waldo, the mother of Defendant Thomas' child. A search of the crime scene at Ms. Waldo's residence in reference to this Destruction of Property resulted in the recovery of one (1) PSD 17 .223/5.56 casing (Item 1 in regard to CCN: 20120626).

On February 17, 2021, Defendant Stevenson was arrested at 3519 Stanton Road, Southeast, with Defendant Phillips present. During the execution of the search warrant, MPD seized a pair pants consistent with those worn by Defendant Stevenson on the day of the homicide, as well as an additional handgun (One (1) black/tan P80 model PF940C 9mm semiautomatic pistol (CS2-1)).

**B. Firearms & Toolmark Report by Chris Monturo[9]**

Chris Monturo examined the firearms evidence in this case the results of which appear in reports dated June 17, 2021, and March 24, 2023.[10] *See* Exhibits 1 and 1a.

As it relates to cartridge casings and firearms recovered from the homicide crime scene and defendants' flight path, Mr. Monturo concluded as follows:

- The Smith & Wesson M&P caliber 9mm semiautomatic pistol (P4-1) recovered from the SPO fired the three (3) SPEER 9mm Luger casings (CS1-1 through CS1-3)

- The Omni Hybrid semiautomatic rifle (CS3-1) fired the PSD 17 .223/5.56 cartridge casings recovered from the homicide crime scene (CS1-4 through CS1-8 and CS1-11 through CS1-39)

- The black and silver Taurus G2c 9mm firearm recovered from Troy Williams fired the RP 9mm Luger cartridge casing (MV1-4) recovered from the rear driver side floor under the mat of the victims' car

- The Omni Hybrid semiautomatic rifle (CS3-1) fired the bullet fragment recovered from the homicide crime scene in the parking lot near Stanton Road (CS1-40) and multiple bullet fragments recovered from the victims' vehicle (MV1-1 through MV1-3, MV1-7, MV1-9, and MV1-11).

[9] Mr. Monturo's *CV* is attached as Exhibit 1b.

[10] The March 23, 2023 report describes the results of analysis and comparison of a cartridge casing from the August 2020 Destruction of Property.

- As it relates to additional bullet fragments recovered from inside the victim vehicle (MV1-10-1 and MV1-8), Mr. Monturo concluded that these fragments shared similar class characteristics with test fired bullets from the Omni Hybrid semiautomatic rifle (rifling widths and type), he concluded that the results were inconclusive due to damage to the fragments from impact and lack of corresponding individual characteristics. Therefore, the Omni Hybrid semiautomatic rifle cannot be identified or eliminated as having fired these bullet fragments.

- Metal bullet fragment MV1-10-2 recovered from the interior of the victims' car, specifically the rear driver's side door panel had no value for comparison or identification purposes.

- Analysis was also conducted of a bullet recovered from the front passenger floor of Defendant Thomas' green Volvo (MV3-9). Mr. Monturo concluded that the Smith & Wesson M&P 9mm (P4-1)(SPO's gun) fired that bullet (MV3-9).

As it relates to evidence recovered from the decedent at autopsy, Mr. Monturo concluded the following:

- The Omni Hybrid semiautomatic rifle (CS3-1) fired bullet fragment A1-6 recovered from the decedent's right arm.

- Metal fragment A1-3 had no value for purposes of comparison or identification.

As it relates to the PSD 17 .223/5.56 cartridge casing recovered in a separate incident on August 23, 2020, Mr. Monturo concluded that it was fired by the Omni Hybrid rifle.

Each of these conclusions was verified by Calissa Carper, a qualified firearms examiner.[11]

### C. Testimony by Chris Monturo

Here, firearms examiner Monturo will opine based on his training and experience and the degree of agreement of random characteristics observed under the comparison microscope. He will **not** use unqualified terms such as "match." He will **not** state his expert opinion with any level of statistical certainty, much less 100% or absolute certainty. He will **not** render his opinion "to the exclusion of all other firearms" or use the phrase "to a reasonable degree of scientific

[11] Ms. Carper's *CV* is attached as Exhibit 1c. These verifications were not blind.

certainty." This is consistent with the Uniform Language for Testimony and Reports for the Forensic Firearms/Toolmarks Discipline – Pattern Matching Examination ("DOJ ULTR") (Govt. Ex. 3a.), which defines "source identification" as "an examiner's conclusion that 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 insufficient disagreement of individual characteristics to conclude they originated from different sources." DOJ ULTR, Section III. The DOJ ULTR precludes examiners from associating a casing to a firearm "to the exclusion of all other sources," from asserting a "numerical degree of probability" without appropriate data, or from using the expression "reasonable degree of scientific certainty." *Id.*

Specifically, Mr. Monturo will testify that where he made a source conclusion that is his opinion that the observed class characteristics and corresponding random characteristics provide extremely strong support for the proposition that the two toolmarks -- i.e., the firearm and bullet/cartridge case – originated from the same source and extremely weak support for the proposition that the two toolmarks originated from different sources. He will state that it is his opinion, based on his examination of class and random characteristics,[12] research supporting the ability to use random characteristics to tell one firearm from another, research indicating that firearms and toolmark examiners generally do so with low error rates, and verification of his conclusions in this case, that the most reasonable explanation is that this firearm fired these cartridge casings.[13] This language

[12] Mr. Monturo may testify regarding his knowledge of subclass characteristics, that his identification is not based on subclass characteristics, and the basis for such conclusion.

[13] Mr. Monturo may use the photographs taken of the submitted evidence in this case and may discuss consecutively manufactured firearms studies (including any he has completed) and other research (generally or specifically).

represents the current understanding of the forensic science of firearms and toolmarks, it is qualified, and informs the jury of the uncertainty inherent in all forensic examination of evidence. The language largely originates from the definition of source identification from the DOJ ULTR and is very similar to the language developed by the Organization of Scientific Area Committees for Forensic Science, Firearms and Toolmark Subcommittee. The Subcommittee developed this language over the course of several years and it involved work and input from the OSAC Subcommittee membership, which include firearms toolmark examiners, statisticians, research scientists, and other forensic pattern disciplines (e.g. fingerprint examiners). The language is currently a proposed standard and is being reviewed by the Academy Standards Board.[14]

**ARGUMENT**

In *Daubert*, the Supreme Court provided a non-exhaustive list of factors to consider when evaluating the admissibility of expert testimony under Rule 702, including whether a theory or technique (1) has been or can be tested; (2) has a known or potential rate of error; (3) has been subjected to peer review and publication; (4) has standards controlling the techniques operation; and (5) enjoys acceptance within the relevant scientific community. *Daubert*, 509 U.S. 579, at 593-94. Rule 702 also allows for expert testimony where the expert "reliably applied the principles and methods to the facts of the case." Fed. R. Evid. 702(d). Rule 702 embodies a more liberal standard of admissibility for expert opinions than did *Frye*. *United States v. Williams*, 506 F.3d 151, 161-62 (2d Cir. 2007). In exercising its gatekeeping function, courts must keep in mind the Supreme Court's admonition that "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." *Daubert*, 509 U.S. at 596.

[14] As a result this language may change.

Courts have routinely admitted firearms evidence under *Daubert* and Rule 702. *State v. Romero*, 341 P.3d 493, 498 (Ariz. App. Div. 2 2014) ("Several federal district courts have held that firearms identification testimony is sufficiently reliable under Daubert and Federal Rule 702."); *United States v. Hicks*, 389 F.3d 514, 526 (5th Cir. 2004) (noting that defendant was unable to point the court to a single case in any circuit showing that the methodology was unreliable); *United States v. Santiago*, 199 F. Supp.2d 101, 111 (S.D.N.Y. 2002) ("The Court has not found a single case . . . that would suggest that the entire field of ballistics identification is unreliable."). Many federal courts have held extensive hearings before admitting firearms and toolmark identification evidence. *See e.g., United States v. Montiero*, 407 F. Supp. 2d 351, 355 (D.Mass. 2006) (six-day hearing); *United States v. Diaz*, 05-CR-00167, 2007 WL 485967, at *1 (N.D. Cal. Feb. 12, 2007) (four-day hearing); *United States v. Taylor*, 663 F. Supp. 2d 1170, 1171 (D.N.M. 2009) (two-day hearing); *United States v. Otero*, 849 F. Supp.2d 425, 429 (D.N.J. 2012) (three-day hearing).

While the issuance of the PCAST Report created a flurry of firearms litigation, it did not significantly alter the legal landscape:

> While no federal court (at least to the Court's knowledge) has found the AFTE method to be unreliable under Daubert, several have placed limitations on the manner in which the expert is allowed to testify. The general consensus is that firearm examiners should not testify that their conclusions are infallible or not subject to any rate of error, nor should they arbitrarily give a statistical probability for the accuracy of their conclusions. Several courts have also prohibited a firearm examiner from asserting that a particular bullet or shell casing could only have been discharged from a particular gun to the exclusion of all other guns in the world. These restrictions are in accord with guidelines issued by the Department of Justice for its own federal firearm examiners which went into effect in January 2019. But it is also important to note that the courts that imposed limitations on firearm and toolmark expert testimony were the exception rather than the rule. Many courts have continued to allow unfettered testimony from firearm examiners who have utilized the AFTE method.

*United States v. Romero-Lobato*, 379 F. Supp. 3d 1111, 1117 (D. Nev. 2019) (citations omitted, emphasis added). *See also United States v. Brown*, 973 F.3d 667, 704 (7th Cir. 2020) (no abuse of discretion in trial court's refusal to adhere to PCAST recommendations and its finding that firearms and toolmark identification is tested, subjected to peer review and publication, has an overall low error rate (single digits), and is generally accepted in the specified scientific field); *United States v. Rhodes*, 2023 WL 196174 (D. Or. Jan. 17, 2023); *United States v. Chavez*, ___ F.Supp.3d __, 15-CR-00285, 2021 WL 5882466 (N.D. Cal. Dec. 13, 2021) (denying request to limit prosecutions' ballistics evidence where defense submitted an affidavit from Dr. Scurich and Mr. Faigman); *United States v. Harris*, 502 F.Supp.3d 28, 38 (D.D.C. 2020) (admitting firearms and toolmark expert testimony with the limitations identified in the DOJ ULTR); *United States v. Hunt*, 464 F. Supp. 3d 1252, 1259-62 (W.D. Okl. 2020) (admitting testimony following AFTE theory of identification and indicating DOJ ULTR is reasonable guidance scope of testimony); *United States v. Johnson*, 2019 WL 1130258, at *21-22 (S.D.N.Y. March 11, 2019) ("In the vast majority of cases in which courts have limited the opinions a firearms examiner may offer, the limitation has addressed whether the firearms examiner can state his or her opinion to a specific degree of scientific certainty. Often these limitations are imposed because of judicial or defense counsel concern that the firearms examiner intends to offer an opinion with absolute or 100% certainty.") (citations omitted); *United States v. White*, 17-CR-611, 2018 WL 4565140 (S.D.N.Y. 2018) (finding no hearing necessary where the admissibility of firearms expert testimony has been repeatedly recognized by federal courts and noting the expert may not quantify or give an opinion to the exclusion of all firearms); *United States v. Gregory Chester, et. al.*, No 13-CR-00774, slip. op. at 2 (N.D. Ill. Oct. 7, 2016) ("In short, the PCAST report does not undermine the general reliability of firearm toolmark analysis or require exclusion of the proffered opinions in this case.

Questions about the strength of the inferences to be drawn from the analysis of the examiners presented by the government may be addressed on cross-examination.") (Govt. Ex. 17); *People v. Lozano-Membrano, et.al.*, No. 1501755, Oral Ruling, p.21 (Sup. Ct. Cal. July 16, 2020) (declining to limit testimony to class characteristics in a case where defendant's submitted an affidavit from Mr. Faigman and ruling that "[t]he expert may testify as to exclusions and inclusions, but must identify the limitations of her opinion that, one, it does not exclude all firearms, two, that it is not presented as a scientific certainty, and three, that they will give no numerical or statistical calculation") (Govt Ex. 18)[15]; *Commonwealth v. Hernandez*, SUCR2014-10417 * 5, slip. op. (Super. Ct. Mass. Dec. 21, 2016) (denying defendant's motion to preclude firearms and toolmark identification evidence based on PCAST) (Govt. Ex. 19); *Commonwealth v. Legore*, SUCR 2015-10363, slip op. at 2 (Superior Court Mass., Nov. 17, 2016) ("After a non-evidentiary hearing and argument, and upon review of the PCAST report (and in particular, pages 104-114), there is no basis to disturb settled law permitting a properly qualified firearms expert from offering opinion evidence under [*Daubert/Lanigan*] relating to a comparison and match between a bullet recovered from the alleged victim, and a bullet test-fired from a firearm allegedly associated with the defendant.") (Govt. Ex. 20).[16]

[15] The ruling and pleadings in Lozano-Membrano addressed a separate case *People v. Azcona*, 58 Cal. App. 5th 504, 510 (2020), *as modified* (Jan. 11, 2021). In that case, the court determined the defendant had failed to establish that firearm and toolmark identification was no longer generally accepted in the applicable scientific community. *Id.* at 512-13. The court further found, however, that it had erred by allowing the expert to testify that the matching marks on the relevant projectiles are "much more than can ever happen by random chance," and therefore the projectiles came from the same gun, "to the practical exclusion of all other guns." *Id.* at 513-14. The court went on to say that "[s]uch a purportedly infallible conclusion is a leap too far from what the underlying method allowed. There was support for the opinion that the projectiles likely came from the same gun, perhaps more likely than not, but there was no basis to present it as a scientific certainty." *Id.* The Azcona court only described the 2008 and 2009 NAS Reports, and 2016 PCAST Reports in support of its determination.

[16] *But see State of Illinois v. Ricky Winfield,* 15-cr-14066 (Cook County January 11, 2023) *(*precluding firearms and toolmark testimony under *Frye*) (transcript available upon request). In precluding the firearm and toolmark testimony, the court specifically adopted the defense brief. Mr. Weller testified in this

Consistent with the weight of authority and based on the extensive testing, peer-review, low levels of error, and general acceptance throughout the world, this Court should admit the expert testimony of Mr. Monturo with the qualifications proposed by the government described above. This is consistent with the law in the District of Columbia, the vast majority of case law in the United States, and the science establishing the reliability of the discipline.

### A. D.C. Court of Appeals Case Law Does Not Require the Limitations Requested by the Defendant.

The defendant requests that the Court exclude Mr. Monturo's testimony in its entirety, or alternatively limit it to "cannot exclude." Def. Mtn. at 11. Neither exclusion, nor the particular language requested, is required by *United States v. Gardner*, 140 A.3d 1172 (D.C. 2016), *United States v. Williams*, 210 A.3d 734 (D.C. 2019) (*Williams II*),[17] or the science. Neither *Gardner* nor

---

hearing. The court found he was not credible, specifically adopting the defense brief in this regard and stating he mischaracterized the studies in the field. The government asserts that the Court should ignore this in its entirety. Certainly the government disagrees with many of the arguments asserted by the *Winfield* court, many of which it addresses in this Opposition. That is the nature of litigation. Aside from that, however, the ruling regarding Mr. Weller lacks specificity. The overall opinion appears biased and considers irrelevant information (such as allegedly false convictions resulting allegedly from firearms and toolmark evidence). Additionally, most of the studies cited in Mr. Weller's January 12, 2022 declaration are attached for confirmation of the information in his declarations. For completeness sake, the government has attached as Exhibit 21, many of the other studies listed in Mr. Weller's declaration but not specifically discussed herein.

[17] Recently, in *Gordon v. United States*, No. 18-CF-1319, 2022 WL 16985005, at *16 (D.C. Nov. 17, 2022), the D.C. Court of Appeals briefly addressed the language of firearms and toolmarks expert conclusions as it relates to answering the question whether cartridge casings or bullets were fired from the same firearm, as opposed to whether a specific, recovered firearm fired one or more cartridges or bullets. Although not directly on point, it is instructive insofar as it appears to state, at least under the circumstances the particular case, that a conclusion that a certain number of cartridge casings were fired from the same gun, when accompanied by the statement that the expert could not "conclusively" provide such opinion, is consistent with *Williams II.* In *Gordon*, the firearms examiner, Christopher Coleman, "discussed the markings he had observed on the six shell casings discovered at the murder scene and two bullets recovered from [the decedent's] body and Ms. Morris's apartment." *Id.* Coleman then explained that "the six casings "most likely" were fired from some type of Glock semiautomatic pistol," that "the two bullets were "consistent" with a Glock, but he could not exclude another type of gun, or say conclusively that they were fired from the same gun." *Id.* "On re-direct, Mr. Coleman testified that "all six cartridge cases were fired in the same gun."" Based on a plain error review, the *Gordon* Court upheld this testimony stating it did not run afoul of *Williams II*. It stated, "*Williams* expressly limited its holding to the precise issue of an examiner providing "unqualified testimony," and did not reach the related issue of whether an expert using toolmark

*Williams II* had the benefit of any information related to Firearms and Toolmark Identification beyond the 2008 and 2009 NAS Reports and PCAST, which were either limited in their examination of the field and are now outdated.[18] More importantly, neither *Gardner* nor *Williams II* had the benefit of studies in the field, or any testimony or affidavits from well-respected members from the field or broader the scientific community. Neither of those cases had the benefit of post-PCAST studies that confirm the reliability of the field. This Court accordingly need not engage in an in-depth examination of the rulings in *Gardner* or *Williams II*, and notably neither opinion required the limitation requested by the defendant. Rather, each simply precluded the kind of unqualified opinion that the government does not intend to offer in this case.[19] Indeed, *Williams II* stated, "[W]e do not question the admissibility of the firearms and toolmark examiner's

---

analysis may link a specific bullet to a specific gun if he does not "do so with absolute or 100% certainty." Id. (citing *Williams II*, pp. 740-41). The Gordon Court then noted that "Coleman neither provided unqualified testimony nor matched a specific bullet *to a specific gun.* Rather, Mr. Coleman was careful to qualify his opinion, and only opined on the fact that the six cartridges were most likely fired from a similar type of *unspecified* gun." *Id.* "Accordingly, the trial court did not err – let alone plainly err – by failing to s*ua sponte* strike Mr. Coleman's testimony." *Id.*

[18] PCAST, which is a policy paper, suffered from numerous flaws that undermine its value. Additionally, as PCAST itself notes, its framework for determining what it calls "scientific validity" is not required for admissibility. PCAST Report at 4. However, because it serves as the primary basis in support of the defendant's request for exclusion or limitation, for purposes of this Opposition, and without conceding the validity or applicability of PCAST or its framework, the government accepts its framework for evaluation of firearms and toolmarks and summarizes the subsequent firearm and toolmark studies that have satisfied the criteria for foundational validity as defined by PCAST.

[19] Specifically, the D.C.C.A. in *Williams II* confronted a record in which the examiner had testified, *inter alia*, "[t]hese three bullets were identified as being fired out of Exhibit No. 58. And it doesn't matter how many firearms Hi[-]Point made. Those markings are unique to that gun and that gun only." *Williams*, 210 A.3d at 738. Similarly, in *Gardner*, the Court's confronted a record where "the prosecutor specifically asked Mr. Watkins, 'Just to be clear, sir, your—your scientific—your opinion here is Government Exhibit Number 18, the bullet, [was] fired from Government 71[,] or was it consistent with being fired from Government Exhibit 71?" Mr. Watkins replied, "It was identified as having been fired from Government Exhibit 71." *Gardner*, 140 A.3d at 1182. The Court of Appeals concluded only that the examiner could not "give an unqualified opinion about the source of the bullet" and further held "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.

testimony generally." 210 A.3d at 743. *Williams II* went on to state,

> Following *Gardner*, we repeat that it is error to allow an examiner to provide this kind of unqualified opinion testimony, but we do not foreclose the possibility that the necessary data will exist at some point in the future to provide a foundation for opinion testimony that unqualifiedly connects a specific bullet to a specific gun.

*Id.*

The scientific information provided herein and at any hearing in this case provides such a foundation. Even adopting the PCAST framework relied on by *Gardner* and *Williams II*, post-PCAST studies establish what PCAST termed "foundational validity." That said, the government does not intend to admit *unqualified* firearm and toolmark identification testimony. Rather, in ensuring Mr. Monturo's testimony complies with the restrictions set out in the DOJ ULTR and in using the conclusion language described above, the government asserts it is in compliance with *Gardner* and *Williams II*.

The government's proposed language is qualified. It recognizes that the conclusion is not to 100% certainty. It does not purport to exclude all other firearms in the world. The probative value of the language is consistent with the scientific research in the field that establishes the reliability of the methodology and ability of qualified examiners to use class and individual/random characteristics to distinguish one firearm from another with low rates of error. It is consistent with conclusion language proposed by the OSAC Firearms and Toolmarks Subcommittee, which was agreed to after significant work by toolmark examiners and experts in other fields, and thus has a basis in the field's language, research, and philosophy.

On contrast, "cannot exclude" extremely underrepresents the probative value of an examiner's identification conclusion given the extensive research regarding the ability to use class and random characteristics to distinguish one firearm from another, the body of research indicating that examiners make such conclusions with low rates of error, and the verification of the results in this case by a second examiner. "Cannot exclude" typically reports a class characteristic inclusion without more. Limiting conclusions to

class characteristics would make the field extremely unreliable as discussed in Mr. Weller's supplemental declaration. *See* 12/9/22 Weller Decl. (Govt. Ex. 2a.) This language does not account for, or report reasonable conclusions based on consistency in class and random characteristics observed by a firearms examiner on the evidence and reference samples in the case. Thus, it would be an incomplete and misleading statement. Whereas a class characteristic inclusion narrows the pool of alternative firearms down somewhat, potentially to 100s or 1000s of firearms that could have left the class marks on the evidence, random characteristics narrow down that number to substantially far fewer. An expert conclusion based on class and random characteristics should reflect the scientific research supporting that such consistency reliably indicates that no more than a small number of other firearms could have left such marks.

Even if the Court were to decree that "cannot exclude" accounted for determinations based on class and random characteristics, this conclusion lacks meaning and does not accurately express the evidentiary weight of the examiner's conclusions. "Consistent with" suffers from the same flaws. "Cannot be excluded" is a double negative, meaning the same as "included." Without additional testimony, this conveys that potentially thousands of firearms or more could also have left the toolmarks, which is inconsistent with the scientific research. "Cannot be excluded" and "consistent with" are similarly unhelpful in that they provide little to no information regarding evidentiary weight, unless the examiner explains to the jury how likely, or not, it is that some other firearm could have made the toolmarks. This would require testimony regarding the ability to use random characteristics to accurately distinguish one firearm from another and to determine that a particular firearm fired particular bullets/cartridge cases.

As such, the government requests that the Court allow Mr. Monturo to testify to the conclusions described above which reasonably and accurately convey the probative value of the conclusion consistent with the scientific research in the field that establishes the reliability of the methodology and ability of qualified examiners to use class and individual/random characteristics to distinguish one firearm from another with low rates of error.

## B. Firearms and Toolmark Identification Satisfies Reliability Under Rule 702

### 1. Testability

Firearm and toolmark identification has been and continues to be tested and found reliable. Testability focuses on "whether the expert's theory can be challenged in some objective sense, or whether it is instead simply a subjective, conclusory approach that cannot reasonably be assessed for reliability." *See* Fed. R. Evid. 702 Advisory Committee's Note to 2000 Amendment. "[V]irtually every court that has evaluated the admissibility of firearms and toolmark identification has found the AFTE method to be testable and that the method has been repeatedly tested." *Tibbs*, 2019 WL 4359486, at *7 (listing cases); *see also Harris*, 502 F. Supp. 3d at 37 (stating that "[a] number of courts have examined this factor in depth to conclude that firearm toolmark identification can be tested and reproduced[,]" and compiling citations including *Tibbs*). "Scientific methodology today is based on generating hypotheses and testing them to see if they can be falsified; indeed, this methodology is what distinguishes science from other fields of human inquiry." *Daubert*, 509 U.S. at 593 (citation omitted).

There is a long and continuous history of firearms and toolmarks research. Since the "phenomenon behind Firearms and Toolmark Examination, namely that firearms can impart microscopic toolmarks on fired ammunition components, and that those toolmarks can be used for the purposes of source attribution (i.e., identification) and elimination was first documented over 100 years ago[,] . . . the profession has been engaged in observation, documentation and testing of firearms examination related topics." *See* 1/12/22 Weller Decl. ¶ B1. Such foundational research appears in textbooks[20] and peer reviewed journal articles. *Id.* Articles often referred to as "Review

[20] There are several textbooks published over a large span of time cited in Mr. Weller's declaration, including RONALD NICHOLS, FIREARM AND TOOLMARK IDENTIFICATION: THE SCIENTIFIC RELIABILITY OF THE FORENSIC SCIENCE DISCIPLINE 1-159 (2018). This text "explains how past and recent research provide strong support for the science of firearm and toolmark examination." 1/12/22 Weller Decl. ¶ B1.

Articles" serve as "encyclopedia-like sources of research that summarize the state of a topic or discipline." *Id.* Attached are three Review Articles related to Firearms and Toolmark Identification, one written by Ronald Nichols and published in the Journal of Forensic Sciences in May 2007,[21] which lists 65 references, the second by Erwin J.A.T. Mattijssen and published in Forensic Science International: Synergy in 2020,[22] with 189 references, addressing research occurring between 2016 and the end of 2018, and the third also by Erwin J.A.T. Mattijssen and published in Forensic Science International: Synergy in 2023,[23] addressing research and advances published from 2019 until and including 2021. (Govt. Exs. 4a, 4b, and 4d respectively). Summaries of foundational research also appear on the AFTE website, which lists over 100 citations with summaries of each article.[24] "These articles, which only represent a fraction of the total body of research, show the profession has published research papers that span over half a century." *See* 1/12/22 Weller Decl. ¶ B1.

Testing in the discipline has establish that consistency in class and random characteristics may be used to distinguish one firearm from another with little to no random matches. That is, the forensic science has established and continues to test toolmark variability.[25] Some of these studies

---

[21] Ronald G. Nichols, *Defending the Scientific Foundations of the Firearms and Tool Mark Identification Discipline: Responding to Recent Challenges*, 52 J. FORENSIC SCIENCES 586, 586–94 (2007) (Govt. Ex. 4a.)

[22] Erwin J.A.T. Mattijssen, *Interpol review of forensic firearm examination 2016-2019*, 6 FOR. SCI INT. SYNERGY 389, 389–403 (2020) (Govt. Ex. 4b.)

[23] Erwin J.A.T. Mattijssen, *Interpol review of forensic firearm examination 2019-2022*, 2 FOR. SCI INT. SYNERGY 100305 (2020) https://doi.org/10.1016/j.fsisyn.2022.100305 (Govt. Ex. 4d.)

[24] AFTE Website, https://afte.org/resources/swggun-ark/testability-of-the-scientific-principle, last accessed 5/8/2023.

[25] Indeed, that firearms can impart individual characteristics/random marks to fired ammunition that can be used to distinguish ammunition fired from one firearm from ammunition fired from another is well established in the field and not challenged by the defendant. Indeed, such a challenge would be extremely difficult to mount as the discriminative capacity of individual marks is largely uncontested:

are briefly discussed in Todd Weller's January 12, 2022 Declaration, in paragraphs B1 through B4. *See* 1/12/22 Weller Decl. Here, we briefly discuss some recent studies that use computer algorithms to compare class and random marks on bullets and cartridge casings, because these studies report large numbers of intercomparisons and thus establish variability with no random matches. Mr. Weller summarizes this point in his declaration:

> First, the computer algorithms used in these studies can be tasked with performing a large number of inter-comparisons. *This type of research rigorously tests the hypothesis of whether same-source and different-source toolmarks can be differentiated*. Second: 3D-based research does not rely on human examiner judgment. Therefore, we can separate examiner performance from the question of toolmark variability.

1/12/22 Weller Decl. ¶B4.

The National Institute of Standards and Technology (NIST) has done substantial work in this area as discussed in paragraph B5, including a 2007 study in which NIST studied 22 firearms

---

> [T]he potential discriminative capacity of traces on elements of spent ammunition, is largely uncontested, due to what is generally known about the large variability of design features introduced during manufacturing processes of firearms, as well as the various mechanisms and phenomena that lead to configurations of surface features that leave highly variable traces on elements of fired ammunition (e.g., Bonfanti and DeKinder, 1999). We will not pursue this topic any further here other than by repeating that features on elements of fired ammunition (bullets and cartridge cases) have intrinsic probative value, conditional on being correctly assessed and interpreted by a proficient and knowledgeable examiner.

*See* Drs. Alex Biedermann, Bruce Budowle, Christophe Champod, *Forensic feature-comparison as applied to firearms examination: evidential value of findings and expert performance characteristics*, April 28, 2022, p.17-18 (Govt. Ex. 15) (hereinafter Biedermann et. al.). Since the "phenomenon behind Firearms and Toolmark Examination, namely that firearms can impart microscopic toolmarks on fired ammunition components, and that those toolmarks can be used for the purposes of source attribution (i.e., identification) and elimination was first documented over 100 years ago[,] . . . the profession has been engaged in observation, documentation and testing of firearms examination related topics." *See* January 12, 2022 Declaration of Todd Weller, ¶ B1 (hereinafter 1/12/22 Weller Decl.) (Govt. Ex. 2).

Variability is not to be confused with uniqueness. The relevant question for the Court is not: Are toolmarks unique? *See* Bunch et al. "At some level all physical objects are unique. Rather, the relevant question is, can a trained human or machine reliably distinguish between toolmarks made by one tool versus toolmarks made by other tools. These consideration lead to the disciplines first basic proposition, a directly testable claim that includes examiners as integral to the science: Class and microscopic marks imparted to objects by different tools will rarely if ever display similarities/correspondence sufficient to lead a qualified firearm-toolmark examiner to conclude that the objects were marked by the same tool. That is, a qualified examiner will rarely if ever commit a false-positive error (misidentification) (Bunch 2008)" *Bunch et al*.

and 178 casings and found an extremely small false identification rate and no report of random matches. *See* Weller 1/12/22 Decl. ¶B5.[26] NIST also completed a study using bullets from the Hamby error rate study, which consisted of 35 bullets fired through 10 consecutively manufactured Ruger barrels (expected to have the highest degree of similarity in random marks from one firearm to the next as they come off the manufacturer line). The study involved 46 known-matching (same-source) comparisons and 549 known-nonmatching (different source) comparisons. They were able to correctly identify all known matching bullets, again supporting the variability of toolmarks and no or limited existence of random matches. *Id.*[27]

This research is not limited to NIST. Bachrach et. al. used 3D analysis methods to examine striated toolmarks in which the algorithms conducted over 100,000 comparisons. There was one false exclusion in a set of 450 matching comparisons (0.11% error rate) and three false inclusions (0.03% error rate) in a set of 4500 nonmatch comparisons. *Id.* at ¶B6.[28]

Todd Weller and several other scientists outside of the firearms and toolmarks discipline conducted a study using 3D technology involving close to 8000 comparisons.[29] "Despite the presence of subclass toolmarks, there was no overlap of data between the matching (720 comparisons) and non-matching (7290 comparisons) test fires." Weller 1/12/22 Decl. ¶B6. The authors reported, in part, "The data strongly support the hypothesis that for the type of

[26] Vorburger et al Surface Topography Analysis for a Feasibility Assessment of a National Ballistics Imaging Database NISTIR 7362, 2007.

[27] Chen Z et al. Fired bullet signature correlation using the Congruent Matching Profile Segments (CMPS) method For Sci Int 2019 (305) http://dx.doi.org/10.1016/j.forsciint.2019.109964 .

[28] B. Bachrach et al. "A Statistical Validation of the Individuality and Repeatability of Striated Tool Marks: Screwdrivers and Tongue and Groove Pliers*", J. Forensic Science, March 2010 (5'5)2 pp. 348-357.

[29] Weller T, Zheng A, Thompson R, Tulleners F "Confocal Microscopy Analysis of Breech Face Marks on Fired Cartridge Cases from 10 Consecutively Manufactured Pistol Slides" J. Forensic Science, July 2012 57(4) pp. 912-917.

manufacturing processes studied, marks left by the breech face can be used both to distinguish between firearms and to associate fired evidence to a particular firearm." *Id.* In a second paper co-authored by Mr. Weller with several other scientists outside of the firearms and toolmarks discipline, 3D imagery technology performed more than 100,000 comparisons with no false positives. *Id.* at ¶B8.[30] As Mr. Weller stated in his declaration,

> The body of novel 3D-based research provides objective support for the prior century's research and experiential knowledge of firearm and toolmark examiners. The use of 3D toolmark measurements and computer comparison has allowed for a *large number* of matching and non-matching toolmark comparisons using objective mathematical models. The resulting data has resulted in non-practitioner researchers concluding the research "support existing conclusions" and "constitute support of the experiential knowledge" and calculate "extremely small false identification error rates.

1/12/22 Weller Decl. ¶B10. These studies all show that where a firearms examiner identifies correspondence in class and random characteristics the likelihood of a random match is extremely low.

The discipline has also studied subclass markings. Such studies have identified certain manufacturing methods under which subclass markings are more and less likely to occur and provide guidance on how to identify and assess their significance in any particular examination of firearms evidence. *See generally* Subclass Characteristics: From Origin to Evaluation, by Ron Nichols. *See generally* Nichols Subclass Review (Govt. Ex. 4c). This confirms the testability and continued testing of the field and discredits the defendant's unsupported assertion that the field lacks standards to distinguish between subclass and individual/random characteristics. Def. Mtn. p.11.

Firearms and toolmark identification research on these and other topics in the field have been published in numerous peer-reviewed journals and conducted by experts in the field and

[30] Weller T, Brubaker M, Duez P, Lilien R "Introduction and Initial Evaluation of a Novel Three-Dimensional Imaging and Analysis System for Firearm Forensics" AFTE Journal, Fall 2015 47(4), pp. 198-208.

experts in other fields such as physical scientists, statisticians, and computer scientists. Mr. Weller's declaration cites to more than forty scientific studies involving firearms and toolmark identification analysis that were published in journals other than the AFTE Journal, including the following ten scientific journals: 1) Forensic Science International, 2) Journal of Forensic Sciences, 3) Science and Justice, 4) National Institute of Standards and Technology, 5) Surface Topography, 6) Measurement Science and Technology, 7) Scanning, 8) Three Dimensional Imaging, Processing and Applications, 9) The Annals of Applied Statistics, and 10) Journal of Physics. *See* 1/12/22 Weller Decl. at 52-55, Appx. A. These studies alone were authored by numerous scientists, many of whom hold PhDs in a wide range of the applied sciences, including statistics, engineering, quantum chemistry, mathematics, physics, computer science, and physical chemistry. *Id.* Many of these scientists, such as Dr. Max Morris, the former Chair of Statistics at Iowa State University, specialize in experimental design.

2. Error rates

The primary issue typically presented at trial is whether the identification rendered by the firearms examiner is, in fact, correct. That is precisely the question presented in the instant case. As discussed in greater detail in this section, to answer that question, the Court should inquire how often examiners err when making a positive identification. *Harris*, 502 F. Supp. 3d at 434 ("[T]he critical inquiry under this factor is the rate of error in which an examiner makes a false positive identification, as this is the type of error that could lead to a conviction premised on faulty evidence."); *Otero*, 849 F. Supp. 2d at 433 ("Indeed, for the purposes of utilizing toolmark identification in legal proceedings, the critical validation analysis has to be the extent to which false positives occur."). The ***seven*** following studies – discussed in more below – demonstrate a low rate of false positive identifications (ranging from zero to 1.01%):

1) David P. Baldwin et al., *A study of false positive and false negative error rates in cartridge case comparisons*, AMES LABORATORY, USDOE, TECHNICAL REPORT #IS-5207 (2014). (Govt. Ex. 5.) ("Ames I")

2) Mark A. Keisler et al., *Isolated Pairs Research Study*, 50 AFTE J 56-58 (2018); Mark A. Keisler et. al., *Letter to the Editor: Isolated Pairs Research Study*, 50 AFTE J 56-58 (2018) (Govt. Ex. 6a and 6b.) ("Keisler" or "Keisler Study")

3) Stanley J. Bajic, *Report: Validation Study of the Accuracy, Repeatability, and Reproducibility of Firearm Comparisons*, AMES LABORATORY-USDOE TECHNICAL REPORT # ISTR-5220 (2020) ("Ames II") (Govt. Ex. 7). This is the prepublication manuscript. It is being published in pieces. The first piece described the study design. *See Planning, design and logistics of a decision analysis study: The FBI/Ames study involving forensic Firearms Examiners*, Keith L. Monson, Erich D. Smith, Stanley J. Bajic, FOREN. SCI. INTL.: SYNERGY 4, at 2 (2022) ("Monson") (Govt. Ex. 7a). The second is the accuracy portion referred to herein as Monson II. *Accuracy of comparison decisions by forensic firearms examiners*, Keith L. Monson, Erich D. Smith, Stanley J. Bajic, J. FOREN. SCI. 68:86:100 (2023) p.96 (Govt. Ex. 7b.)

4) Large scale research study of cartridge case comparison that is currently undergoing peer review. The study was conducted by Dr. Max Guyll, Stephanie Madon, Dr. Kayla Burd and others. Because it is in pre-publication review, the manuscript is not attached. However the study and its findings are described in the attached declaration by Dr. Guyll. (Govt. Ex. 10.)

5) Pierre Duez et al., *Development and Validation of a Virtual Examination Tool for Firearm Forensics*, 63 J. FORENSIC SCI. 1069, 1069-1084 (2018) (Govt. Ex. 8.) ("Duez" or "Duez Study")

6) Chad Chapnick et. al., *Results of the 3D Virtual Comparison Microscopy Error Rate (VCMER) Study for Firearm Forensics*, J. FORENSIC SCI. at 1 (Oct. 1, 2020) (Govt. Ex. 9.) ("Chapnick" or "Chapnick Study")

7) Neuman M et. al., *Blind testing in firearms: Preliminary results from a blind quality control program*, J. FORENSIC SCI. 2022, pg.1-11 DOI: 10.1111/1556-4029.15031 (Govt. Ex. 11). ("Neuman")

It is important to note that these error rates do not represent the percentage of time any particular examiner would make a false positive, or that the percentage of time a lab would report a false positive.[31]

[31] *See e.g.*, Ames I at 19 ("This finding does not mean that 1% of the time each examiner will make a false-positive error. Nor does it mean that 1% of the time laboratories or agencies would report false positives, since this study did not include standard or existing quality assurance procedures, such as peer review or

The defendant's challenge centers on three outdated policy papers, namely the 2008 and 2009 NAS Reports and the 2016 PCAST Report. The 2016 PCAST Report reviewed literature in the field and criticized the design of most of the studies that at that time validated the accuracy of firearms and toolmark identification. PCAST Report, p.106. Nevertheless, such firearm and toolmark identification has a sound scientific basis.[32] Contrary to conclusions in the PCAST Report, numerous pre-2016 studies demonstrate the validity of firearm and toolmark identification, many of which are contained in the resources cited above and discussed in Mr. Weller's January 12, 2022 declaration.

It is unnecessary, however, for this Court to undertake a detailed examination of these three reports, because even accepting PCAST's concerns, recent studies confirm that the field meets and exceeds the benchmarks for scientific validity set by its critics. PCAST determined that two black-box, sample-to-sample/open studies, employing the use of independent comparisons would be necessary to establish the "foundational validity" of Firearms and Tool Marks Identification, and that one was already in existence, specifically the Ames I study.[33] *See* 2016 PCAST Report at 109-11. In a law review article, Dr. Eric Lander, co-Chair of the PCAST Report reiterated that

---

blind reanalysis) (Govt. Ex. 5); Ames II ("Definitive false positive error rate estimates that take examiner heterogeneity [i.e., errors were clustered among a minority of examiners] into account are 0.66% for bullets and 0.933% for cartridge cases. False negative error rate estimates are 2.87% (bullets) and 1.87% (cartridge cases") (Govt. Ex. 7.) *See also* 1/12/22 Weller Decl. ¶¶ C22 & C29.

[32] Faults with PCAST's criticisms of firearms and toolmarks identification are described in a Department of Justice Report. *See* U.S. Department of Justice Statement on the PCAST Report; Forensic Science in Criminal Courts: Ensuring Scientific Validity of Feature-Comparison Methods (January 2021) (Govt. Ex. 3b.) Mr. Weller's declaration provides further information related to the reliability of the PCAST Report and its criticism of firearm and toolmark identification. *See* 1/12/22 Weller Decl. at 47-51.

[33] As noted in Mr. Weller's declaration, "While [Ames I is] not published in peer reviewed journal, this study has undergone extensive review by both firearms and toolmark examiners as well as the PCAST commission. Given this extensive review by the general scientific community and the lack of any critique, it is unlikely a journal-based peer review would result in any substantial changes." 1/12/22 Weller Decl. Appx. A, n.221.

even under PCAST's standards (which exceed what is required under Rule 702) only one additional black box study was required to establish "foundational validity." *See* 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).

PCAST defined a "black-box study" as an "empirical study that assesses a subjective method by having examiners analyze samples and render opinions about the origin or similarity of such samples." PCAST Report at 48. PCAST further stated, "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" PCAST Report at 49. PCAST stated the following as it relates to open and closed set designs in firearm and toolmark examinations:

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

PCAST Report at 108. PCAST also indicated the study should include what it terms "independent" comparisons, which it defines as "examiners making a series of independent comparison decisions between questioned sample and one or more known samples that may or may not contain the source." *Id.* at. 110 (the second portion of the sentence reiterates the definition of "open"). This study design is referred to in Todd Weller's declaration as sample-to-sample, open, and will be referred to herein as such or simply as "black box study." *See* 1/12/22 Weller Decl. ¶ C21.

The "single well-designed study" referenced by PCAST and Dr. Lander is Ames I. In Ames I, 218 tests were returned. The study calculated a false positive rate of 1.01%, false negative rate of 0.367%, sensitivity of 98.6%, and specificity of 65.2%.[34] *See* Ames I, Table III, p.17 (Govt. Ex. 5.); 1/12/22 Weller Decl. ¶ C22. According to Dr. Lander, if one more study reproduced the results of Ames I, which Dr. Lander described as "well-designed," it would render firearms identification "scientifically valid." *Lander*, 86 FORDHAM L. REV. at 1672 ("A second study would solve this problem."). This is consistent with the position of the PCAST Report, which described Ames I as an "appropriately designed" black box study and explained that one additional, similar study would establish "foundational validity."[35] PCAST Report at 4, 109–111.

To the extent this specific PCAST threshold is required by Rule 702 or *Daubert*, which we do not concede, the field has and continues to respond with additional research and study. The field has responded with additional research and study. In addition to Ames I, there are ***six*** additional studies, five of which are black box (sample-to-sample/open) as defined by PCAST and one of which describes the results of five years of blind proficiency testing, which PCAST stated should be "vigorously pursued," and "could provide valuable insight about the actual error rates in casework." *See* PCAST Report at 59, 97. In short, the firearms and toolmark field has responded to PCAST with more than ample data to establish foundational validity, and more importantly the reliability of the methodology. The ***six*** are briefly described below (and attached).

[34] "Sensitivity is the portion of times examiners reported identifications when the ground truth is same source (examiners correctly reported an identification). Specificity is the portion of times examiners reported eliminations when ground truth is difference source (examiners correctly reported elimination)." 1/12/22 Weller Decl. ¶ C2.

[35] PCAST stated that a showing of foundational validity "requires that a method has been subjected to empirical testing by multiple groups, under conditions appropriate to its intended use. The studies must (1) demonstrate that the method is repeatable and reproducible, and (2) provide valid estimates of the method's accuracy (that is, how often the method reaches an incorrect conclusion) that indicate the method is appropriate to the intended application." PCAST report at 5.

The black box studies employing traditional examination techniques are listed chronologically first ("Keisler," "Ames II," and "Guyll"), followed by black box studies employing 3D computer comparisons ("Chapnick" and "Duez"), followed by the study describing the results five years of blind proficiency testing ("Neuman").

In 2018, the ***Keisler*** Study was published in the AFTE Journal. It adheres to the sample-to-sample/open, black-box requirements set by PCAST. In the Keisler Study, examiners completed 126 tests resulting in 1512 same-source and 1008 different-source comparisons. No false identifications or eliminations were reported. *See* Keisler, et. al, Winter 2018, p.57. The Keisler Study reported a sensitivity of 99.74% and specificity of 79.86%. *Id.*; *see also* 1/12/22 Weller Decl. ¶ C24.

Ames II is the most recent published black box/open set study to report low false positive rates. "In 2012, FBI researchers began to design a decision analysis study that would assess the performance of F/T examiners." Monson at 2. It "assessed the performance of numerous qualified firearms examiners working in accredited laboratories in the United States in terms of overall accuracy (error rate), repeatability, and reproducibility of decisions involving forensic comparisons of simulated firearms evidence (bullets and cartridge cases)." *Id.* at 3. As stated in Monson, the design features included:

- Study participants shall be anonymized, qualified F/T examiners who are AFTE members working in accredited laboratories

- Firearms from three different manufacturers and multiple examples of the chosen models, including those that are deemed relatively difficult to compare. Firearms and ammunition will be selected that tend to produce limited microscopic marks for comparison and no aperture shear, but present subclass characteristics.

- Groups of consecutively manufactured slides and barrels that are collected at intervals throughout the manufacturing life of the single tool used to cut/shape them, to produce highly similar but individual (non-matching) specimens (best known non-match) and maximizing the potential for subclass similarity

- Additional comparison slides and barrels from different production runs (known non-match)

- Extensive firing of each firearm (~500x) to test effects of firing sequence on the reproduction and longevity of individual characteristics, thereby affecting examiner accuracy

- Preparation and distribution of test packets and use of double-blind conditions to conduct comparisons

- An open set design i.e., there may not necessarily be a match for every Q specimen

- To increase the relative test difficulty, an overall proportion of true matches of approximately 33%, but variable among test packets.

- All items in an individual comparison set shall be fired from the same make and model firearm, precluding elimination based on class characteristics

- A break-in period of firing new firearms to normalize marks they produce

- A comparison set consisting of a single Q to be compared to two K specimens, the latter being fired from the same firearm. Providing multiple K specimens minimizes the possibility that a single K did not replicate a toolmark.

- Each set represents an independent comparison unrelated to any other set in the test

- Survey of participants, to include laboratory accreditation, personal certification, years of experience, equipment used in comparison, and laboratory policies on inconclusive and exclusionary decisions

- Discourage collaboration or verification of decisions by a second examiner

- Preclude sharing of results on individual packets and the possibility that participants may infer test design by coding/relabeling the contents before their submission to another examiner or resubmission to the same examiner.

- Pilot testing to evaluate study design

Monson at 4. "The overall false positive error rate was estimated at 0.656% and 0.933% for bullets and cartridge cases, respectively while the rate of false negatives was estimated at 2.87% and 1.87% for bullets and cartridge cases, respectively." Ames II at 2. Sensitivity was 76.6% for bullets and 74.4% for cartridge cases. "These estimates are based on the beta-binomial probability

model and do not depend on an assumption of equal examiner-specific error rates." Ames II at 2. The authors concluded that a majority of errors were produced by a relatively small number of examiners. *Id.* at 73, 77. Overall, as the Court can see, and as the Ames II authors concluded, the accuracy results from this study were consistent with Ames I.[36] *Id.* at 71-72.

Confirming the low error rates of the methodology reported by Ames I, Keisler, and Ames II, is another black-box study by Dr. Max Guyll (and others) that has been accepted for publication (peer-reviewed). Importantly, Dr. Guyll is a relative new-comer to firearms and toolmarks error rate studies. He is an Associate Professor of Psychology with tenure in the School of Social and Behavioral Sciences at Arizona State University. He is also a faculty member of the Law and Behavioral Sciences Initiative, which focuses on issues related to the intersection of psychology and law. Dr. Guyll's research focuses on the effect of human factors in forensic analysis. Dr. Guyll, along with others, completed a large-scale validity study focused on cartridge case comparisons. The results of the study, "[c]ollapsing across firearm model[,]. . . indicated a total of five false-positive errors and one false-negative error, corresponding to an overall false-positive error rate of 0.6% and an overall false-negative rate of 0.1%." *See* Guyll Decl. ¶B. (Govt. Ex. 10.) Overall, the study concluded that the "findings of the research study are consistent with the interpretation that cartridge case comparison is a forensic technique that is characterized by low error rates, and that examiner's conclusive decisions possess strong probative value for the determination of ground truth." *Id.*

Two additional post-PCAST studies using 3D image technology further support the field's

[36] The Ames II authors concluded that with regard to repeatability and reproducibility participating examiners generally exceeded expected agreement. *Ames II* at 77. The data from Ames II are similar to fingerprint examination, which also uses pattern recognition. 1/12/22 Weller Decl. ¶C32 (demonstrating by using table and data from Ulery, B.T., Hicklin, R.A., and J. Buscaglia *Repeatability and reproducibility of decisions by latent fingerprint examiners* PLoS ONE (2012)).

reliability and establish "foundational validity." A 2018 study referred to herein as the Duez Study,[37] included black-box (sample-to-sample/open-set) and white-box components and used Virtual Comparison Microscopy (VCM), which allows for side-by-side comparisons of toolmarks on fire ammunition components using computer software. The empirical data illustrated the low rate of errors made by trained firearm examiners in identifying casings to a particular firearm. In the open-set black-box component, forty-six trained examiners "correctly reported 100% of the identifications (known matches) while reporting no false positives."[38] 1/12/22 Weller Decl. ¶C26.[39] Among trained examiners, sensitivity was 100%, i.e., 276 identifications from 276 true same-source comparisons. *Id.* Specificity was 87%. i.e., 80 eliminations from 92 true different source comparisons. *Id.*

In a 2020 study, referred to herein as the Chapnick Study, firearm examiners used 3D technology that, according to the authors, will "potentially replace the light comparison microscope as the primary instrument used for firearm and toolmark examination."[40] The study involved 76 trained firearms examiners from the United States and Canada and 40 test sets of fired cartridge casings from firearms with a variety of makes, models and calibers. 1/12/22 Weller Decl. ¶ C28. These 76 examiners completed a total of 1184 comparisons. *Id.* The overall error rate for

[37] Pierre Duez et al., *Development and Validation of a Virtual Examination Tool for Firearm Forensics*, 63 J. FORENSIC SCI. 1069, 1069-1084 (2018). Both parts (black box and white box) of this study survived double blind peer-review in the Journal of Forensic Sciences (JFS), demonstrating that the scientific community rejects the notion that only black box design is worthy of consideration. The involvement of two Ph.D. scientists – Drs. Marcus Brubaker and Ryan Lilien – further refute the argument that applied scientists are not involved in the validation of firearms and toolmark identification (Govt. Ex. 8).

[38] A trainee reported two false positive results. *See* Weller. Decl. ¶ C26.

[39] Importantly, the government admitted this study at the *Tibbs* hearing. The import of this study was not appreciated by the *Tibbs* court because it went unmentioned in the *Tibbs* ruling.

[40] Chad Chapnick et. al., *Results of the 3D Virtual Comparison Microscopy Error Rate (VCMER) Study for Firearm Forensics*, J. FORENSIC SCI. at 1 (Oct. 1, 2020) (Govt. Ex. 9.).

this group was 0.253%. *Id.* Three false positives were reported from a total of 693 true different source samples, for a false positive error rate of 0.433%. *Id.* There were no false eliminations reported. *Id.* Sensitivity was 92.2% and specificity was 62.9%. 1/12/22 Weller Decl. ¶ C28.

Although not a black-box study, but equally probative to the reliability of the firearms and toolmarks methodology, a study published in the Journal of Forensic Science in March 2022, details the results of blind, proficiency test case results from the Houston Forensic Science Center between December 2015 and June 2021. Maddisen Neuman, et. al., *Blind testing in firearms: Preliminary results from a blind quality control program*, J. FORENSIC SCI. (March 1, 2022) ("Neuman Study") (Govt. Ex. 11). In blind proficiency testing, mock cases are created to mimic case work. A blind case is assigned to an examiner after which it undergoes microscopic examination and comparison. *Id.* p.964. There were no false positive or false negative results. The results are, in part, as follows:

> Satisfactory results were obtained for all items evaluated, or by the "hard error" definition [25], no hard errors were observed; that is, no identification were declared for true nonmatching pairs, and no eliminations were declared for true matching pairs. The ground truth was compared to the examination results, and the ground truth was obtained 59.7% (n=33) of the comparisons. In 40.3% (n=225) of the comparisons, an inconclusive conclusion was made when the ground truth was either elimination or identification. A ground truth of elimination and comparison result of inconclusive was observed more frequently at 74% (n-106), while the ground truth of identification and comparison result of inconclusive was observed at a rate of 31% (n=119). All items submitted as ground truth insufficient or unsuitable were satisfactorily determined as such. Furthermore, no ground truth submissions of identification or elimination were determined to be unsuitable or insufficient. Inconclusive decisions were only reported for items with a known ground truth of identification or elimination; however, due to the quality of the evidence submitted, inconclusive results were not unexpected.

*Id.* p.968. The study further stated, ". . .[T]hese results indicate firearms examiners routinely reach a correct determination of ground truth identification for cartridge cases and bullets (more sensitivity) but may have] more difficulty discriminating elimination in bullets compared with

cartridge casings (less specificity). *Id.* p.971.

PCAST opined that from a scientific perspective, an acceptable error rate should be less than 5%. PCAST Report, p.152. As shown above, these are far below that. As low as these error rates are, the error rate in casework is likely lower. Many of these studies were designed to be more difficult than casework in several respects. For example, the Ames II study designers chose to use consecutively or closely manufactured firearms and steel ammunition. 1/12/22 Weller Decl. ¶¶ C29 & D2.

Most importantly, these studies overestimate the potential frequency of errors in that they lack a second-level review, an important step to ensure as close to zero errors in casework as possible. Labs employed by the government employ a second examiner who re-examines the results of the first examiner. *See* 1/12/22 Weller Decl. ¶ H4. This is likely to decrease the rate of errors.[41] Additionally, a defendant may elect to have his or her own expert re-examine the evidence to ensure against the unlikely occurrence of a false identification.[42]

Given the low overall error rates, and low false positive error rates associated with the firearms and toolmarks identification methodology, the mechanisms in place in casework to ensure that reports do not issue with erroneous results, and the potential for independent examination by the defendant, this factor weighs in favor of admission of the testimony as proposed by the government.

3. Peer-Review

Studies testing the foundational research of firearms and toolmark identification and

[41] For example, the Ames II data related to reproducibility showed that none of the false identifications were reproduced by a second examiner. 1/12/22 Weller Decl. ¶ C31. "This strongly supports the quality control measures forensic laboratories employ, such as verification. During verification, a second examiner will review the evidence to determine if the same result is obtained as the primary examiner. The Bajic [Ames II] reproducibility data provides evidence that the forensic laboratory practice of verification can catch false positive errors." *Id.*

[42] To date, Defendant Stevenson has not requested independent examination.

examiners' ability to associate a cartridge case/bullet to a particular firearm have been authored by a variety of scientists, most of whom hold PhDs in a wide range of the applied sciences -- including statistics, engineering, quantum chemistry, mathematics, physics, computer science, and physical chemistry – and have undergone peer-review in a variety of scientific journals. *See* 1/12/22 Weller Decl. Appx. B (Non-Exhaustive List of Firearm and Toolmark Research by Individuals Associated with Non-Crime Lab Institutions.) In fact, nearly all of the citations in Mr. Weller's declaration are from peer-reviewed scientific journals.[43] Moreover, the field's research into objective means to quantify and validate firearms and toolmark identifications through the use of 3D topographical imaging, sophisticated computer algorithms, and the implementation of statistical tools through the establishment of firearms databases is conducted in peer-reviewed scientific journals as well.

Given the breadth of peer-reviewed publication, courts have consistently held that this factor weighs in favor of admissibility. *Harris*, F. Supp. 3d at 39.; *Johnson*, 2019 WL 1130258, at *16; *Romero-Lobato*, 2019 WL 2150938, at *5 ("Several published federal decisions have also commented on the AFTE Journal, with all finding that it meets the Daubert peer review element."); *Ashburn*, 88 F.Supp.3d at 246 ("The court finds that the AFTE methodology has been published and subject to peer review, weighing in favor of admission . . ."); *Diaz*, 2007 WL 485967, at *8 ("The fact that the articles submitted to the AFTE Journal are subject to peer review weighs strongly in favor of admission."); *Otero*, 849 F. Supp. 2d at 433 (noting AFTE Journal's formal process for the submission of articles); *Taylor*, 633 F. Supp. 2d at 1176 (finding the peer review factor "clearly weighs in favor of admissibility.").

PCAST and the outlier decisions cited by the defense seek to discard the vast body of peer-

[43] Mr. Weller has served as peer reviewer for both the Journal of Forensic Science and AFTE Journal. Although the process differs, the result is the same: articles are reviewed by subject matter experts who judge the paper on its scientific merit. Mr. Weller has accepted, revised, and rejected papers for both.

review and publication by taking the position that only studies that undergo what they deem to be "meaningful" peer review should be considered, and all other scientific data should be discarded altogether. There are several flaws in this approach. First, and foremost, as discussed at length in Mr. Weller's declaration, several studies have been peer reviewed and published in a variety of scientific journals. For example, both Duez and Chapnick in which Mr. Weller collaborated with Dr. Lilien and other scientists, were published post-PCAST in the Journal of Forensic Sciences. These and other post-PCAST data survived the very type of peer-review that the critics such as Judge Edelman in *Tibbs* deemed "meaningful."

Moreover, the idea that only one particular type of peer review is "meaningful" is simply wrong. Dr. Bruce Budowle is in a unique position to comment on the peer-review process as the most published forensic DNA scientist in the world who sits on the editorial boards of several scientific journals. Dr. Budowle Decl. at 1 (Gov. Ex. 12.) According to Dr. Budowle, the *Tibbs'* opinion's critique of certain peer review processes is out of step with how the scientific community evaluates peer-review. Leaders in the scientific community, including Dr. Budowle, do not embrace the idea that only double-blind peer-review constitutes "meaningful" review; rather all peer review has value. *Id.*, at 2. Having reviewed Dr. Budowle's affidavit, Judge Contreras questioned "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*." *Harris*, 502 F.Supp.3d at 40.

Additionally, *Tibbs* was critical of the AFTE journal because it did not previously employ double-blind peer review (AFTE now employs double-blind peer review). However, there is no consensus in the scientific community that double-blind peer review is the only meaningful kind; rather, there are three commonly used forms of peer review, one of which is the open peer-review

used (at the time) by the AFTE journal. *See* 1/12/22 Weller Decl. ¶¶ F1-F4. As noted by Mr. Weller, Wiley Publishers, which publishes more than 1600 academic and scientific journals, outlines the pros and cons of each type of peer-review and notes that the use of open peer review is growing. *Id.* ¶ F1. Moreover, for many prestigious journals, double-blind peer review is a recent phenomenon. For example, the prestigious journal Nature first instituted double-blind peer-review in 2015. Even now, such review is voluntary for authors, most of whom do not opt for it (from March 2015 to February 2017 only 12% of Nature authors opted for double-blind review). *Id.* ¶ F3. Does this mean that every scientific discipline that published in Nature prior to 2015, or studies in which the authors did not opt for double-blind review, were somehow deficient under *Daubert*? This would require reevaluating, e.g., whether DNA is a double helix. *See* Watson, J. D., & Crick, F.H. C. A Structure for deoxyribose nucleic acid, Nature 171, 737-738 (1953). Dr. Budowle echoes this sentiment: "[E]ach [form of peer-review] has some merit; however, there is no consensus at this time on which approach is superior. Nonetheless, they all serve as part of the peer review process." Dr. Budowle Decl. at 1 (Govt. Ex. 12.)

Finally, the opinion that only double-blind peer-review amounts to a more "meaningful" review is belied by studies that were initially published in the AFTE Journal and subsequently published in the Journal of Forensic Science with no alterations to the design study. *See Harris*, 502 F.Supp.3d at 40.[44] For example, 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 SCI (2019) (Gov. Ex. 13), is an update of a continuing study that has been ongoing for

[44] The Court in *Harris* stated, "Compellingly, the government also refuted the allegation by Judge Edelman in *Tibbs* that the AFTE Journal does not provide 'meaningful' review, by bringing to the Court's attention a study that was initially published in the AFTE Journal, and then was subsequently published in the Journal of Forensic Science with no further alterations."

over twenty years. *Id.* at 551. The study used the consecutive barrel (worst-case scenario) design. *Id.* Although PCAST was critical of this design, the study survived the double-blind peer-review process in JFS. *Id.* The first four installments of the ongoing study consisted of two AFTE Journal publications, a dissertation, and a presentation at an AFTE training seminar. *Id.* at 557. According to PCAST (and *Tibbs*), the type of peer-review involved in the first four rounds of this study would be a basis to neglect the data altogether. However, the fifth installment of the study (which included all prior four installments) survived double-blind peer review at JFS. Notably, the study design remained unchanged from the earlier installment published in the AFTE Journal. *Id.* (referring readers to the earlier studies "for the complete design of the study"). This illustrates two points: 1) the data first published in the AFTE Journal was worthy of double-blind peer review later published in JFS and 2) the scientific community does not subscribe to the PCAST theory that only black box studies are worthy of scientific publication.

This Court should further reject claims from the defendant's, Def. Mtn. p.7, and *Tibbs* that "the AFTE Journal's use of reviewers exclusively from within the field to review articles created for and by other practitioners in the field greatly reduces its value as a scientific publication, especially when considered in conjunction with the general lack of access to the journal for the broader academic and scientific community as well as its use of an open review process." *See Tibbs*, 2019 WL 4359486, at *10. The assumptions underlying this assertion by the defendant and from *Tibbs* are simply inaccurate, rendering the related conclusions unreliable. In contrast to the understanding of the *Tibbs* Court, the AFTE Journal has a free, publicly accessible online searchable index.[45] 1/12/22 Weller Decl. ¶ F5. The AFTE Journal is also listed in the SCOPUS database, "the largest abstract and citation database of peer-reviewed literature: scientific journals,

[45] https://afte.org/afte-journal/searchable-journal-index

books and conference proceedings."[46] *Id.* It is publicly accessible and non-AFTE-members can subscribe to it, and search for, find, and purchase individual articles. AFTE journal subscriptions have been available to non-AFTE members since at least October 1989.[47] 1/12/22 Weller Decl. ¶ F5. The online, searchable index where pdf versions of articles can be purchased by non-members has been available since 2011.[48] *Id.*

In further contrast to the *Tibbs* Court's (mis)understanding, the AFTE Journal is available at academic institutions. *See id.* It may be freely searched on WorldCat.org and search results show the AFTE Journal is available at Cal State Sacramento, Nebraska Wesleyan University, University of Central Oklahoma, Truman State University, Grambling State University, University of Wisconsin-Platteville, Case Western Reserve University, Marshall University, Radford University, University of Toronto, West Virginia University, Stetson University College of Law, George Mason University, George Washington University, the Library of Congress, Syracuse University, Virginia Commonwealth University, Florida International University, John Jay College, The British Library (UK), University of Wolverhampton (UK), Cranfield University (UK), Bibliotheque de l'EPFL (Switzerland), ESR Mt Albert Science Center (New Zealand), ESR

[46] A list of SCOPUS "sources," available at https://www.scopus.com/sources.uri provides a list of "sources", showing the AFTE Journal is included in this database. "Scopus is the largest abstract and citation database of peer-reviewed literature: scientific journals, books and conference proceedings. Delivering a comprehensive overview of the world's research output in the fields of science, technology, medicine, social sciences, and arts and humanities, Scopus features smart tools to track, analyze and visualize research." https://service.elsevier.com/app/answers/detail/a_id/15534/supporthub/scopus/#tips, accessed 1/14/22.

[47] See "Publication Information" from the AFTE Journal, October 1989, Volume 21, No 4: "Members of the Association of Firearm and Tool Mark Examiners receive a subscription to the AFTE Journal as a part of their dues…A subscription to the AFTE Journal is available by contacting the Editor: M James Kreiser, 2168 South 9th Street, Springfield Illinois, 62702, U.S.A. The subscription rate is $50.00 per calendar year."

[48] Clow, C., Letter to the Editor: *Message from the Chair of the AFTE Journal Index Committee*. The online AFTE Journal Index was launched on January 1, 2011 and is currently available at https://afte.org/afte-journal/searchable-journal-index. Individual articles can be purchased for $5.00 each.

Kenepuru Science Center (New Zealand), NT Police, Fire and Emergency Services Library (Australia) and the University of Western Australia. 1/12/22 Weller Decl. ¶ F7.

Defendant wrongly asserts and *Tibbs* wrongly concluded that AFTE and the AFTE Journal is an insular organization comprised solely of examiners in the field, who isolate themselves from the greater scientific community and academics, and who are motivated to confirm their own beliefs, "comparable to talk within congregations of true believers" rather than by a sincere desire to engage in critical scientific review and debate. *See Tibbs* 2019 WL 4359486, at *10. *Tibbs* ruled on this notion, along with the inaccurate assumption that only double-blind peer-review is valuable, to reduce the value of and ignore scientific data published in the AFTE Journal. *See id*. Given the faulty factual basis for this conclusion, this Court should disregard the defendant's arguments in this regard and *Tibbs* and consider the many AFTE Journal articles, along with relevant articles in other publications in determining admission of the testimony at issue. *See* 1/12/22 Weller Decl. ¶¶ F2-F11, Appx. A. Ignoring such articles and the data therein simply because of the journal of publication is not good science. It is more akin to ignoring data simply because it conflicts with the desired outcome or hypothesis. The Court should not rely on conclusions reached through the application of such unscientific principles.

4. Standards

Standards and controls for the firearms and toolmarks profession are published and maintained from several sources. AFTE has published the following standards for professional guidance and use:

- AFTE Training manual: 166-page document outlining all steps a new trainee should undertake prior to starting casework.
- AFTE Technical Procedures Manual: 116-page document providing technical procedures for typical examinations that may occur in firearms and toolmark identification laboratories.

- AFTE Glossary: 244-page document providing the profession with standardized terminology and definitions.[49]

- AFTE Theory of Identification *(see supra pp. 7-8)*:[50] First published in 1992, the Theory of Identification distills the essence of firearms identification into several paragraphs. The Theory of Identification describes the basis for an identification as well as the limitations of the profession (i.e., that identification/same-source conclusions are not absolute).

1/12/22 Weller Decl. ¶ G1.

The AFTE Theory of Identification sets the standard for how an examiner learns the amount of microscopic (tool)mark agreement expected in same source (i.e. known match) versus difference source (i.e. known non-match) samples. The AFTE Theory specifically directs the profession to seek out and use worst case scenario ("best known non-match") samples as part of training. Examiners learn to recognize the amount of agreement associated with worst case scenarios, or "best known non-matches," by examining fired bullets and cartridge cases collected from sequentially manufactured firearms where the similarity in (tool)marks from one firearm (or tool) to the next is at its greatest ("worst case scenario") because the firearms are taken off the manufacture line one after another.

While this decision making standard includes some subjectivity – which is true of decisions in almost all forensic disciplines including DNA analysis – it is bound by these concepts defined by the AFTE Theory of Identification upon which all examiners are trained. Criticisms that the standard is so subjective that different examiners trained and working in different labs at different locations will have wildly different decision thresholds for satisfaction of AFTE Theory of Identification are undermined by common experience and the scientific research. First, if this criticism had any merit one would expect to see more errors in the field. Yet, errors are extremely

[49] The AFTE Training Manual, Technical Procedures Manual and Glossary are too large to be attached to this pleading. They are available for download at www.AFTE.org.

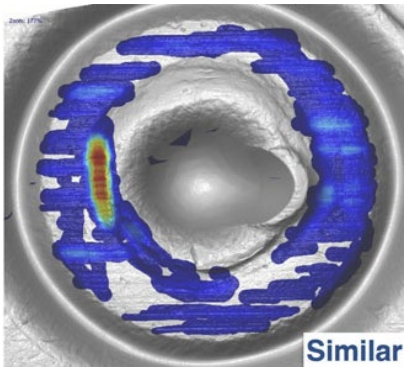
[50] Committee for the Advancement of the Science of Firearm & Toolmark Identification, *Theory of Identification as it Relates to Toolmarks: Revised*, 43 AFTE JOURNAL at 287 (2011).

rare in casework. If the discipline were as unreliable as some of its critics suggest, defense experts would routinely produce results different from government experts. Notably, the defendant does not point to a single case where an independent defense examination led to a different conclusion than the one sponsored by the United States.

Second, the field's scientific research supports the hypothesis that qualified examiners generally apply the consistent decision thresholds. The consistently low error rates in both closed and open-set/black box error rate studies support this. If qualified examiners had significantly different decision thresholds, the false positive and false negative error rates in these studies would be significantly higher. Additionally, two white box studies support this proposition. White-box studies are performed "to understand the factors that affect examiners' decisions." PCAST Report, p.9. The National Institute of Science and Technology (NIST) Glossary defines white box testing as, "[A] method of testing software that tests internal structures or workings of an application as opposed to its functionality (i.e., black-box testing). *See* https://csrc.nist.gov/glossary/term/white_box_testing .

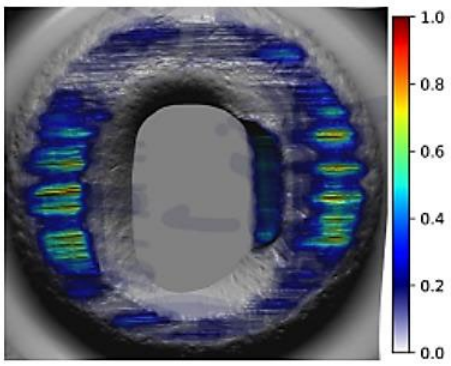
Duez contained a white-box aspect (in addition to a black box aspect discussed above). Duez establishes that even examiners from numerous different laboratories (15), each working independently, mostly use the same amount and same location of microscopic marks when concluding identification. *See* 1/12/22 Weller Decl. ¶ C27; Duez et al. at 1069-1084. The software allowed each examiner to independently annotate the areas where he or she identified significant agreement. Weller Decl. ¶ C27. At the completion of all tests, each annotation was overlaid to compare where each examiner found significant agreement. *Id.* Following is a heatmap of breach face markings (from Duez) showing the combination of these annotations of significant agreement.

The closer to red, the closer to 100% of examiners that marked that area. Id. The red and orange areas indicate a high degree of correspondence of marks by examiners. Id.




Thus, not only did Duez reaffirm that trained examiners could accurately identify casings to a specific firearm, but it also showed that trained firearms examiners focus on the same location and amount of marks to make a conclusion, and that alternative forms of study design are valuable to the advancement of the field.

Similarly, Chapnick, in which 40 different tests sets were independently compared by approximately 30 different examiners, generated similarity heat maps for each set of comparisons. *See* 1/12/22 Weller Decl. ¶ C28; *see generally* Chapnick, et al. An example is below (Figure 3):




Set 1

The Chapnick study authors concluded, in part: "Examiners had a high amount of agreement with regard to the areas useful for identification and elimination as well as those areas which should be avoided for definitive source attribution. It is worth reiterating that examiners worked independently and that the described annotation map patterns emerged when these independent

submissions were combined. This consistency reinforces the fact that examiners typically agree on the toolmarks most important and most reliable for reaching source conclusions." Weller Decl. ¶ C28; Chapnick, et al, p.13.

Another source of standardized guidelines was established by the Scientific Working Group for Firearms and Toolmarks (SWGGUN). *Id.* SWGGUN was a committee of firearms and toolmark examiners whose responsibility was to publish guidelines for the firearm and toolmark community. It has been replaced by the Organization of Scientific Area Committees (OSAC) for which Mr. Weller currently serves as the Vice Chair, and previously served as the Chair. SWGGUN guidelines are on the OSAC website.[51] OSAC is in the process of revising and writing discipline-specific standards. *Id.* It has published seven standards to the OSAC registry and six draft standards and is currently drafting additional best practices and standards. *Id.* National Institute for Standards and Technology also provides metrological standards. It currently offers a standard bullet and cartridge case and is working to produce a new set of reference standards for use in emerging 3D technology.[52]

These standards for the field of firearms and toolmarks, the efficacy of which is supported by the research in the field, satisfies this factor.[53]

---

[51] https://www.nist.gov/topics/forensic-science/firearms-and-toolmarks-subcommittee

[52] Stocker, M. et al., *Addressing Challenges in Quality Assurance of 3D Topography Measurements for Firearm and Toolmark Identification,* 50 AFTE JOURNAL 104-111 (2018).

[53] Courts have come to different conclusions regarding this factor. *Compare Johnson*, 2019 WL 1130258 at *17-18 (standards weighs in favor of admissibility) *with Harris*, 502 F. Supp. 3d at 41 ("While a close call, the Court finds that the lack of objective standards ultimately means this factor cannot be met.") However, disagreement between courts on this one factor has not prevented the vast majority of other courts from finding that firearms and toolmark identification satisfies *Daubert*. This is not surprising because the lack of objective standards or subjectivity of a methodology has little to do with whether the methodology has been tested, found to have low error rates in ground truth studies, and been subjected to peer review and scientific publication. *Harris*, 502 F. Supp. 3d at 42 ("It should be noted, however, that even if this factor cannot be met, a partially subjective methodology is not inherently unreliable, or an immediate bar to admissibility. Rule 702 'does not impose a requirement that the expert must reach a conclusion via an

5. General Acceptance

A "reliability assessment does not require, although it does permit, explicit identification of a relevant scientific community and an express determination of a particular degree of acceptance within that community." *Daubert*, 509 U.S. at 594. "Widespread acceptance can be an important factor in ruling particular evidence admissible, and a known technique which has been able to attract only minimal support within the community may properly be viewed with skepticism." *Id*. The overwhelming majority of published opinion evaluating the admissibility of firearms and toolmark evidence has found that the AFTE Theory of Identification enjoys general acceptance in the relevant community and that such acceptance weighs in favor of admissibility. *See, e.g., Romero-Lobato*, 379 F. Supp. 3d at 1122; *Johnson*, 2019 U.S. Dist. LEXIS 39590, at *58, 2019 WL 1130258, at * 19; *Johnson*, 2015 U.S. Dist. LEXIS 111921, at *11, 2015 WL 5012949, at *4; *Ashburn*, 88 F. Supp. 3d at 247; *Wrensford*, 2014 U.S. Dist. LEXIS 102446, at *45-46, 2014 WL 3715036, at *14; *Taylor*, 663 F. Supp. 2d at 1178.[54]

This makes sense as firearms and toolmark identification is practiced worldwide in government and independent forensic laboratories and studied worldwide by toolmark and scientists with expertise in multiple fields including applied physics, computer science, chemistry, metallurgy, and statistics. Specifically, firearms and toolmark identification is practiced by accredited laboratories throughout United States and throughout the world, including England

objective set of criteria or that he be able to quantify his opinion with a statistical probability.'") (*quoting Romero-Lobato*, 379 F. Supp. 3d at 1120).

[54] *Tibbs* recognized that all the opinions evaluating the admissibility of firearms and toolmarks found that the AFTE Theory of Identification enjoy general acceptance in the relevant community. *Tibbs* declined to follow the weight of legal authority stating that "courts must not confine the relevant scientific community to the specific group of practitioners dedicated to the validity of the theory—in other words, to those whose professional standing and financial livelihoods depend on the challenged discipline." *U.S. v. Tibbs*, No. 2016-CF1-19431, 2019 WL 4359486, at *21.

(Scotland Yard), New Zealand, Canada, South Africa, Australia, Germany, Sweden, Greece, Turkey, China, Mexico, Singapore, Malaysia, Belgium, Netherlands, and Denmark. 1/12/22 Weller Decl. ¶ H1. In the United States alone, there are 233 laboratories with firearm and toolmark accreditation certificates, all of which routinely identify bullets or casings to a particular firearm. *Id.* The firearms units do not exist in a vacuum, but rather are part of a greater scientific accreditation umbrella; virtually all of these accredited firearms units function within a larger forensic laboratory offering a multitude of accredited scientific units, *e.g.*, chemistry, DNA, latent fingerprint identification, etc. In the greater Washington, D.C. metropolitan area alone, FBI and ATF maintain accredited firearms and toolmark units, along with a variety of accredited forensic disciplines that support local and federal investigations. Notably, PhD scientists in various scientific disciplines are in charge of many of these laboratories. For example, the Houston Forensic Science Center (HFSC), an independent forensic lab often identified by defense counsel and some experts as an extremely progressive and reliable lab, is led by CEO Peter Stout, a PhD toxicologist. HFSC's Firearms Unit issues reports with the range of conclusions consistent with that listed in the AFTE Theory of Identification. *See* HFSC Website, https://records.hfscdiscovery.org/Published/Firearms%20Section%20Range%20of%20Conclusions%20Effective%2010-3-2022.pdf (last visited 5/7/2023). In addition, as discussed *supra*, scientists who design, conduct, and publish validation studies in the area of firearms and toolmark identification accept the AFTE method of identification, as do leading scientific working groups such as the OSAC, which is comprised of a variety of applied scientists. *See* 1/12/22 Weller Decl. Appx. B-Non-Exhaustive List of Firearm and Toolmark Research by Individuals Associated with Non-Crime Lab Institution. Indeed, many of the papers discussed in detail were the result of collaborations between firearms and toolmarks examiners and experts in many other fields.

6. Chris Monturo reliably applied the Firearms and Toolmark Identification methodology.

Mr. Monturo applied the AFTE Theory of Identification reliably. Mr. Monturo is an extremely experienced and qualified firearm and toolmark examiner who has practiced for more than 20 years. From 1995 to 2019 he was an examiner with the Miami Valley Regional Crime Laboratory. From 2019 to present he has worked at the Hamilton County Coroner's Office /Crime Laboratory. Both of those agencies are accredited. Mr. Monturo also runs Precision Forensic Testing, where he conducts examinations as an independent contractor. He is certified in Firearm Identification and Toolmark Identification by AFTE. To obtain such certification an examiner must pass written and practical examinations in the area of certification and recertify every five years. *See* Govt. Exs. 1d & 1e, Mr. Monturo's and Ms. Carper's AFTE Certifications.[55] In addition, Mr. Monturo has always passed his proficiency tests. He has published a book, Forensic Firearm Examination (Academic Press, 2016). He has published several articles in the AFTE Journal. *See* Govt. Ex. 1b (Monturo CV.) In addition, he stays current in his field by undergoing annual training. *See id.* In sum, he can be trusted to apply the principles of firearm and toolmark identification reliably.

In this particular case, he conducted the examination consistent with accepted practices and procedures in the field. As indicated above, for more than 20 years he has worked and continues to work for accredited labs. He applies the procedures used in such labs to his work as an independent contractor. Although the organization under which conducts his independent work is not accredited, that does not render Mr. Monturo's work unreliable. Neither Rule 702 nor *Daubert* require accreditation. Importantly, accreditation, although not without meaning, is not the talisman

[55] See https://afte.org/afte-certification/certification-program for requirements and procedure for certification and re-certification.

of reliability. Rather, Mr. Monturo's training and experience provide a substantial basis to conclude he reliably applied the principles of firearm and toolmark identification in this case.

In addition, Mr. Monturo documented his work in this case and all of his conclusions were verified by another highly qualified examiner, Calissa Carper. Here, the evidence was sent by FedEx with a signature required, ensuring knowledge of the evidence's location at all times.[56] Mr. Monturo then received the evidence himself and inventoried it. He photographed the packaging of all the items, establishing he received them intact and sealed. He then inventoried the evidence, taking note of every item. At his private lab, he maintains two safes to which only he has the code. This is where he stored the evidence in this case to ensure its security. He took notes to document his examination. All of his conclusions were then verified by Ms. Carper.

Ms. Carper is also a very experienced examiner. She has been engaged in firearms and tool examination since 2009 (training from 2009 to 2011, with independent casework starting thereafter). She has worked in such capacity at the accredited West Virginia State Police Forensic Laboratory since 2009. (*See* Govt. Ex. 1c, Ms. Carper's *CV*.) She has always passed her proficiency tests. She is AFTE certified in the area of Firearm Evidence Examination and Identification (since 2019). (*See* Govt. Ex. 1e.) She was an instructor at the ATF's National Firearms Examiner Academy in 2021, among other teaching duties. Ms. Carper has also served as an ANAB[57] assessor.

In sum, the identification conclusions and other results to which Mr. Monturo will testify in this case have been examined by two highly qualified, AFTE certified examiners, who have worked and continue to work in accredited labs for a combined total of more than 30 years. They

[56] FedEx is a common method for transferring evidence from one lab or person to another lab or person.

[57] ANAB provides accreditation services. See https://anab.ansi.org/ (last visited January 8, 2021).

have extensive knowledge of the procedures used at accredited labs and acted consistently in conducting the examinations in this case. Thus, the Court can have a degree of confidence that the examiners in this case reliably applied methodology underlying firearms and toolmarks identifications.

### C. Defendant's Claims Are Without Merit

#### 1. 2008 Ballistic Imaging Report & 2009 NAS Report are Outdated

The 2008 Ballistic Imaging Report is outdated by over a decade. More to the point, Dr. John E. Rolph, the Chair of the Ballistic Imaging Report, put to rest (shortly after the report's issuance) any effort by litigants to use the report as a statement against the scientific validity of firearms and toolmark identification:

> The statement in the Report that the "validity of the fundamental assumptions of uniqueness and reproducibility of firearm-related toolmarks has not been fully demonstrated" (Report at 3-22) was not made in the context of assessing the admissibility of firearms-related evidence. Indeed, the Report states clearly that *"this study is neither a verdict on the uniqueness of firearm-related toolmarks generally nor an assessment of the validity of firearms identification as a discipline.* Our charge is to focus on 'the uniqueness of ballistic images' –that is, on the uniqueness and reproducibility of the markings (toolmarks) left on cartridge cases and bullets as they are recorded or measured by various technologies ...." Report at 1-5 (emphasis in the original)….
>
> The Committee's cautionary statement [about not casting conclusions in terms of absolute certainty, to the exclusion of all other firearms, or implying a zero error rate] is not a commentary on the admissibility of firearm-related toolmark evidence. In the Committee's view, "statements on toolmark matches (including legal testimony) should be supported by the work that was done in the laboratory, by the notes and documentation made by examiners, and by proficiency testing or established error rates for individual examiners in the field and in that particular laboratory." Report at 3-23 to 3-24.

See Sworn Statement of Dr. John E. Rolph (Gov. Ex. 14) ¶¶ 6 & 10 (quoting portions of the 2008 NAS Report). Courts have relied upon Dr. Rolph in placing the 2008 NAS Report in proper context. *See United States v. Casey*, 928 F. Supp. 2d 397, 400 (D.P.R. 2013) (Dr. Rolph's

statements greatly undermine the portions of the 2008 NAS report upon which defendant … relies]."); *State v. Langlois*, 2 N.E.3d 936, 945 (Ohio Ct. App. 2013) ("[T]he 2008 NRC report addressed the issue of establishing a nationwide database for the computer imaging of bullets. The report's primary focus was not firearms identification, comparative ballistics, or tool mark analysis."); *United States v. Taylor*, 663 F. Supp. 2d 1170, at 1175-76, 1179-80 (D.N.M.) (holding firearms evidence admissible after considering, *inter alia*, the 2008 Ballistic Imaging Report); *United States v. Otero*, 849 F. Supp. 2d 425, 430, 438 (D. N.J.) (same).[58]

Similarly, the 2009 NAS Report is only one year less outdated than the 2008 Ballistic Imaging Report. Like the earlier NAS report, defense attorneys did not gain any traction in attempts to use the report as a basis to exclude firearms and toolmark identification testimony in the manner defendant is attempting to do so here. *See United States v. Ashburn*, 88 F. Supp. 3d 239, 274 (E.D.N.Y. 2015) (holding that firearms and toolmark identification is a proper subject of expert testimony under Rule 702 and *Daubert* after considering the 2009 NAS Report); *State v. Romero*, 341 P.3d 498, 498 (Ariz. App. Div. 2 2014) (same); *United States v. Casey*, 928 F. Supp. 2d 397, 400 (D.P.R. 2013) (ballistics evidence admissible under Rule 702 and *Daubert*); *State v. Langlois*, 2 N.E.3d 936, 950 (Ohio Ct. App. 2013) (firearm and toolmark identification satisfies the test for reliability under Rule 702); *Otero*, 849 F. Supp. 2d at 438 (D.N.J. 2012) (essential foundations for the admission of expert testimony under Rule 702 established by the government); *United States v. Willock*, 696 F. Supp. 2d 536, 568 (D. Md. 2010); *Taylor*, 663 F. Supp. 2d at 1180 ("The evidence before the Court indicates that when a bullet is fired from a gun, the gun will impart

[58] In *Tibbs*, Judge Edelman concluded that the 2008 NAS Report "directly addressed the sufficiency of the published studies purporting to show a low error rate in the field of firearm and toolmark identification." Judge Edelman's holding failed to consider the plain and unambiguous sworn statement by Dr. Rolph, refuting his finding, as well as judges who cited Dr. Rolph's statement in coming to the opposite conclusion. (*See* Govt. Ex. 14, Dr. Rolph's Statement.)

to the bullet a set of markings that is, at least to some degree unique to that gun. The evidence further indicates that an experienced firearms examiner can make observations of those markings, using a method that has been peer-reviewed, that allow him, in some cases, to form an opinion that a particular bullet was or was not fired from a particular gun. The court therefore concludes that the firearms identification testimony is admissible under Rule 702 and *Daubert*."); *State v. Lee*, 2017 WL 1494012, *10 (4th Cir. Apr. 26, 2017) ("[E]ven after publication of the NAS Report, courts have addressed, in detail, the reliability of [firearms and toolmark identification] testimony and ruled it admissible, although to varying degrees of specificity."); *Spears v. Ryan*, 2016 WL 6699681, *5 (D. Ariz. Nov. 15, 2016) ("[T]he NAS Report would have had no effect on the admissibility of the toolmarks evidence in this case."); *Napier v. Commonwealth*, 2014 WL 3973113, *9 (Ky. Ct. App., Aug. 15, 2014) (It was not the purpose of the 2009 NAS Report to opine on the long-established admissibility of toolmark and firearms testimony in criminal prosecutions and there was no error in taking judicial notice of scientific reliability of ballistic analysis under *Daubert*); *United States v. Sebbern*, 2012 WL 5989813, at * 8 (E.D.N.Y. 2012) (no need for a *Daubert* hearing before admitting ballistics evidence); *United States v. Cerna*, 2010 WL 3448528, at *5 (N.D. Cal. Sept. 1, 2010) (the NAS report "does not necessitate exclusion of expert [ballistics] testimony").

For these reasons, Judge Contreras concluded that the 2008 Ballistic Imaging Report and 2009 NAS Report are "outdated by over a decade due to intervening scientific studies," and therefore "have been repeatedly rejected by courts as a proper basis to exclude firearm and toolmark testimony." *Harris*, 502 F.Supp.3d at 35.

2. <u>The case law relied upon by defendant is outdated and/or flawed.</u>

This Court should not follow the cases cited by the defendant addressing the admissibility

of firearm and toolmark identification testimony under *Daubert* or *Frye*. None of these opinions is binding on the Court (outside of *Williams I and II*). All but one[59] were issued prior to PCAST and publication of black box error rate studies Keisler, Ames II, Duez, Chapnick, Guyll, and Neuman, which confirm the low error rates of methodology. The government has addressed *Tibbs* above as it relates to the relevant factors under Rule 702. Importantly, *Tibbs* did not have the benefit of evidence related to the Ames II, Guyll, or Neuman studies. It is also worth noting that *Tibbs* took great pains to discount numerous pre-PCAST studies, Ames I (which even PCAST considered valid), and the Keisler Study (the study specifically addressed in the written *Tibbs* opinion). Substituting its own judgment for that of the scientific community (toolmark examination and beyond), *Tibbs* effectively disregarded all of the research supporting a contrary view. For example, *Tibbs* eliminated studies from consideration because of what the Court deemed inadequate instructions to examiners. [60] *See Tibbs*, 2019 WL4359486, at *13. It eliminated studies because the method for selecting participating firearms examiners "does not provide the clearest indication of the accuracy of the conclusions that would be reached by average toolmark examiners." *Id.* at *14. Furthermore, *Tibbs* inaccurately characterized AFTE and the AFTE Journal to discount all of the studies published therein. *See supra* Section B3. If one is looking to find fault in a particular study, one will, because no study, irrespective of design, is perfect. *See generally* Biederman et al. p.7; Dr. Max Morris Declarations[61] (Govt. Ex. 16). Despite this, *Tibbs*

[59] *United States v. Jovon Medley*, No. PWG 17-242 (D. Md. April 23, 2018). As of this submission, the link in defendant Stevenson's motion, and a search of Westlaw, have failed to produce this opinion. In any event, as indicated in defendant's Stevenson's brief the testimony was not excluded in its entirety. Additionally, given the date, it could have considered Ames II, Duez, Chapnick, Guyll, or Neuman.

[60] The *Tibbs* court listed the Stroman, Keisler, and Smith et al (2016) Studies. 4359486, *13.

[61] Dr. Max Morris, a PhD Statistician, discusses study design flaws alleged in many firearm and toolmark studies. In a broad brush description of complex material, he discusses how the alleged flaws are not fatal to the probative value of the studies.

discounted studies based on aspects it found less than ideal rather than look at the universe of research as a whole. In sum, given the incompleteness and basic misunderstandings of such opinions, this Court should decline to follow them.

## Conclusion

For the foregoing reason, defendant's motion to exclude or limit the proposed firearms and toolmark identification testimony should be denied.

Respectively submitted,

MATTHEW M. GRAVES
UNITED STATES ATTORNEY

/s/ *Sharon Donovan*
Sharon Donovan
Miles Janssen
Zachary Horton
Assistant United States Attorneys
USAO-DC, 601 D Street, Northwest
Washington, D.C. 20530

## Certificate of Service

A copy of this filing was served via the CaseFileXpress system upon counsel for Defendant Stevenson, Elizabeth Weller, this 8th day of May 2023

/s/ *Sharon Donovan*
Assistant United States Attorney

.