IN THE UNITED STATES DISTRICT COURT
FOR THE DISTRICT OF THE VIRGIN ISLANDS

UNITED STATES OF AMERICA	:	CRIMINAL ACTION
	:	
v.	:	NO. 3:19-79
	:	
IVAN JAMES, KAI JAMES, JOH	:	
WILLIAMS, MALACHI BENJAMIN,	:	
JAHKIEBO JOSEPH, ARIEL	:	
PETERSEN	:	

MEMORANDUM with FINDINGS

KEARNEY, J.   **January 16, 2024**

Our Grand Jury charges Ariel Petersen, Ivan James, and Joh Williams with multiple firearms offenses among other offenses.[1] Our Grand Jury specifies the illegal use of a Glock 40 caliber pistol, an Olympic Arms MFR AR rifle, and a Romarm WASR-10 AK rifle[2]

Defendants challenge the reliability of expert testimony the United States intends to adduce at trial from proffered experts in firearms toolmarks examination: Kevin Rippman, Major Wells, and Reynold De Souza.[3] Defendants request we limit the scope of the experts' testimony regarding firearms examination and comparison. They claim the science of firearm toolmark examination does not meet the reliability standards defined by the Supreme Court in *Daubert*.[4] The Defendants request we (1) exclude testimony casings and bullets came from a particular firearm at issue in the case or from the same firearm, (2) limit the examiner to a discussion of class characteristics,[5] and (3) exclude comparison photographs purporting to show a match.[6]

Defendants argue the science of firearm toolmark identification does not meet *Daubert's* reliability requirement. Defendants argue three reports show firearms toolmark analysis is unreliable: the 2008 National Research Council (NRC) Ballistic Imaging Report (2008 NRC Report), the 2009 NRC Forensics Report (2009 NRC Report), and the 2016 President's Council

of Advisors on Science and Technology (PCAST) Report (PCAST Report). Defendants argue these reports show (1) no general acceptance of the theory of firearms examinations, (2) no evidence examiners regularly reach accurate conclusions, (3) lack of testing protocols and controls for human error, (4) lack of peer review and publication, and (5) lack of a known error rate. Defendants cite cases limiting the testimony of firearms examiners on *Daubert* grounds.

The United States responds the science of firearm and toolmark examination has been accepted for over 100 years, is reliable, and the cases cited by Defendants are outliers. The United States cites to additional studies conducted since the ones presented by Defendants which demonstrate the science of firearms identification (1) can be and has been tested, (2) has been subject to peer review, (3) has a low error rate (citing studies with errors rates mostly of 0% and as high as 1.6%), (4) is controlled by a set of standards (noting all three labs offering evidence in this case operate under written protocols dictating each step in the testing process), and (5) enjoys general acceptance in the scientific community.[7] The United States cites cases upholding the use of firearms examiners as experts and permitting their testimony.

We allow the challenged expert opinion except the witnesses may not opine to a 100% certainty and otherwise subject to cross-examination.

I.   Adduced evidence.

We granted Defendants' request for an evidentiary hearing to gather evidence on the reliability of the proffered opinions. The United States presented testimony from Todd Weller, Kevin Rippman, and Major Wells. The United States also introduced multiple studies purporting to show the reliability of firearms toolmark examination.

a.   Todd Weller's testimony

Todd Weller swore the American Federation of Toolmark Examiners (AFTE) certified him as a firearm and toolmark examiner since 2008 and he has years of academic and specialty training in the discipline. While he did not test the charged firearms in this case and is not on the United States witness list for trial, Mr. Weller convincingly discussed the theory, practice, and process of firearms toolmark examination. He took issue with the Defendants' reliance on the 2008 and 2009 NRC reports stating the reports themselves concede they are not aimed at admissibility at trial of the science of firearm toolmark examination. He discussed class, subclass, and individual characteristics and how each characteristic is used in firearm examination. In sum, each level of characteristic helps sort the evidence presented to an examiner, with the ability to include or exclude a particular firearm as the source of the toolmarks on shell casing or bullets becoming greater as an examiner proceeds through an analysis of each level of characteristic.

Mr. Weller detailed the process of firearms toolmark examination. First, an examiner evaluates if the evidence provided is testable, *i.e.* are the items presented a firearm and a fired bullet or spent casing. Second, the examiner sorts by class characteristics, *i.e.* a firearm chambered in 9 mm and a 9 mm bullet. Third, the examiner uses a comparison microscope to identify toolmarks left on the bullet or casing during firing. The examiner's results are then subject to verification by another examiner independently repeating this process.

To evaluate the firearm and its toolmarks, examiners compare known matches and non-matches to the evidentiary casings and bullets. Examiners create known matches by test firing the subject firearm and collecting the spent bullets and casings. Known non-matches are produced by test firing a different firearm, but of the same class or subclass to the subject firearm. The possible conclusions are (1) identification, meaning the same source fired the known match and evidentiary

material, (2) elimination, meaning a different source fired the known match and evidentiary materials, (3) inconclusive, meaning the examiner cannot reach a conclusion, and (4) unsuitable, meaning the materials presented are not able to be tested.

Mr. Weller discussed the industry accepted standards and theory of firearm toolmark examination as promulgated by the AFTE, the professional certifying association for firearm toolmark examiners. The AFTE requires an examiner determine there is "sufficient agreement" before declaring a same source identification. "Agreement is significant when the 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."[8]

Mr. Weller described the testability of the science of firearms toolmark examination and its error rate. He swore "black box studies" presented the best studies of testability and error rate. These studies do not test how the process works, but instead focus on if the process achieves reliable, accurate, reproducible results. This is done by using known matches and non-matches and checking an examiner's results against the known truth. The PCAS report cited by Defendants endorses this study approach. In studies using this approach, which most closely replicates the type of examination done in field work, the error rate was consistently less than 2%. All the studies cited are published in peer reviewed journals with at least three of them published after the three studies cited by Defendants. Mr. Weller also noted the studies intentionally use consecutively manufactured firearms to create a "worst case scenario" where consecutively manufactured items will have the least variation in their toolmarks as compared to firearms produced at different times. Some studies also use steel bullets, which also increases the difficulty of identifying the source of toolmarks because steel does not mark as well as brass, which is the more common material

examiners see during field work. Together, these extra hurdles in the academic studies show their resulting error rate is likely higher than the error rate in practical field work.

Mr. Weller detailed the prevalence of firearm and toolmark identification. He noted 200 labs conducting firearms toolmark analysis in the United States and identified at least eight other countries where it is also common practice. He stated independent accrediting boards certify all labs. The AFTE serves this function in the United States.

Mr. Weller admitted on cross examination there is no ultimately determinative objective criteria for declaring an identification. While the indicators, particularly the striae left on the casing or bullet, can be numbered and have led to certain numerically based methods of matching, such as Consecutively Matching Striations (CMS), there is no magic number of matching marks that conclusively demonstrates a particular firearm as the source of a casing or bullet. Mr. Weller uses the Pattern Matching method, which uses objective criteria, such as the striae, to inform the examiners final, subjective sufficient agreement determination.

b.   Kevin Rippman's testimony.

Kevin Rippman is a Bureau of Alcohol, Tobacco, Firearms, and Explosives (ATF) firearm examiner based in the ATF lab in Atlanta, Georgia. Mr. Rippman examined the Glock 40 caliber pistol described in Counts Nine, Twelve, and Thirteen.[9] He testified he applied the identical process described by Mr. Weller. He convincingly described the process of checking for marks, comparing known and unknown samples to the evidentiary materials, and how another ATF examiner independently double-checked his work. He knew of the CMS method, but like Mr. Weller he used Pattern Matching to make his determinations. Pattern Matching is accepted as a methodology in the field of firearm toolmark examination. On cross, he agreed Pattern Matching

has a subjective component, but asserted the subjective final conclusion was based on the objective findings of the examiner in microscopically studying the samples.

c.   Major Wells's testimony.

Major Wells is an ATF firearm and toolmark examiner currently based at the ATF lab is Washington, D.C., but previously based in the ATF Atlanta lab. Mr. Wells analyzed two firearms in this case, an Olympic Arms MFR AR rifle identified in Counts Nine, Twelve, and Thirteen, and a Romarm WASR-10 AK rifle identified in Count Nine. He swore he used the same process described by Mr. Weller and used by Mr. Rippman. He confirmed after his analysis a second examiner performed an independent check of his work and arrived at the same conclusion. Mr. Wells stated the examiners use a comparison microscope which allows them to see two items (bullets or casings) at the same time under the same light and magnification. He swore he used the CMS method as part of his examination but did not write down the number of matching striae. He based his opinions on the "standard," or Pattern Matching, method. He confirmed another examiner could, and did, duplicate his results despite having no record of his observed number of matching striae. He agreed there is no such things as a "100% match, but that is not how we do it." He instead looks for "sufficient agreement," which he defined as an evidentiary sample which exceeds the correspondence of the best-known non-match and exhibited sufficient agreement with the best-known match. His definition aligns with the AFTE definition cited above.

II.   Analysis

Federal Rule of Evidence 702 provides "[a] witness who is qualified as an expert by knowledge, skill, experience, training, or education may testify in the form of an opinion or otherwise if: (a) the expert's scientific, technical, or other specialized knowledge will help the trier of fact to understand the evidence or to determine a fact in issue; (b) the testimony is based on

sufficient facts or data; (c) the testimony is the produce of reliable principles and methods; and (d)

the expert has reliably applied the principles and methods to the facts of the case."

The Supreme Court in *Daubert* requires expert testimony meet three tests: qualification,

reliability, and fit.[10] The party offering the expert must prove each of these requirements by a

preponderance of the evidence.[11] We are focused here on the second prong of reliability. Our Court

of Appeals instructs us to apply an eight-part balancing test to determine reliability:

> (1) whether a method consists of a testable hypothesis;
> (2) whether the method has been subjected to peer review;
> (3) the known or potential rate of error;
> (4) the existence and maintenance of standards controlling the technique's operation;
> (5) whether the method is generally accepted;
> (6) the relationship of the technique to methods which have been established to be reliable;
> (7) the qualifications of the expert witness testifying based on the methodology; and
> (8) the non-judicial uses to which the method has been put.[12]

To be reliable, an expert's assertion need not be verifiable as an objective certainty, but it

must have been derived by a verified scientific process.[13] Absolute certainty is not required.

Rather, "[a]s long as an expert's scientific testimony rests upon 'good grounds, based upon what

is known,' it should be tested by the adversary process – competing expert testimony and active

cross-examination – rather than excluded from jurors' scrutiny for fear that they will not grasp its

complexities or satisfactorily weigh its inadequacies."[14]

Defendants rely on judges' analyses not going as far as Defendants may wish them to, and

acceptance of firearms and toolmark analysis remains widespread.[15] For example, in *United States*

v. Adams, Judge Mosman limited the proposed United States expert to observational statements

only.[16] Judge Mosman prohibited opinion testimony about shell cases matching the firearm at

issue. The expert in *Adams,* however, fumbled badly during his explanation of the process and

procedures used to test the firearms. Judge Mosman limited the ruling to the facts of that case: "I want to be clear that my ruling . . . is limited by the testimony before me during the hearings held in this case. It is not an indictment of forensic evidence or toolmark comparison writ large."[17]

We see similar issues in *United States v. Shipp* (cited by Defendants) where Judge Garaufis denied the defendants' motion for a *Daubert* hearing on firearms toolmark examiner expert and based on the submissions of the parties ruled the United States expert could "testify that the toolmarks on the recovered bullet fragment and shell casing are consistent with having been fired from the recovered firearm, and that the recovered firearm cannot be excluded as the source of the recovered bullet fragment and shell casing.[18] However, [the expert] may not testify, to any degree of certainty, that the recovered firearm is the source of the recovered bullet fragment or the recovered shell casing."[19] Judge Garaufis acknowledged this limitation was "in line with, albeit slightly more restrictive than, limitations that other federal district courts have placed on toolmark analysis testimony."[20] Thus while sometimes imposing outer limits on testimony, our survey of caselaw confirms the vast majority of our colleagues find firearm toolmark analysis reliable and ultimately admissible under Rule 702 and *Daubert*. [21]

III.   Findings

We find, after evaluating witness credibility, reviewing the parties' submissions, and studying the evidence adduced at the hearing, the science of firearm toolmark examination as presented by the United States in this case passes muster under *Daubert's* reliability prong:

1.   The witnesses convincingly demonstrated firearms toolmark examination science consists of a testable hypothesis which posits the toolmarks on an item of known origin can be compared to the toolmarks on an item of unknown origin to arrive at a conclusion about its provenance.

2.   Firearm toolmark analysis is subject to peer review. The theory of the science is reviewed through academically rigorous studies published in peer-reviewed journals. Individual conclusions in a lab are subject to peer review through secondary, independent testing to corroborate or dispel the original test.

3.   The error rate found in academic studies which most closely mirror field work is low. The error rate in field work is likely lower than the error rate in academic studies due to the intentional hurdles placed in the way of the examiners in the studies and the additional safeguards in place in real world labs. While there is no error rate for the field, the studies demonstrate an error rate of between 0% and 2%. These rates are low enough to demonstrate reliability mindful they are subject to cross-examination.

4.   The United States adduced evidence demonstrating the field of firearm toolmark examination maintains standards controlling the testing process and technique. The AFTE oversees and certifies these standards.

5.   The use of firearms toolmark examination throughout the United States and the world, and the majority view of our colleagues, demonstrate its general acceptance.

6.   The Pattern Matching technique employed by Messrs. Rippman and Wells has been found through testing to be reliable.

7.   Neither the Defendants nor the United States presented evidence on the non-judicial uses of firearm and toolmark examination, and this factor is neutral.[22]

---

[1] Third Superseding Indictment, ECF No. 555.

[2] *Id*. While these counts only apply to Defendants Ivan James and Williams, a forfeiture allegation for all firearms applies to Defendant Petersen.

---

[3] The United States only plans to elicit an expert opinion from Messrs. Rippman and Wells. It plans to have Mr. DeSouza offer fact testimony as a chain of custody witness and not an expert. ECF No. 664 at 3. Mr. Rippman and Mr. Wells, but not Mr. DeSouza testified during our evidentiary hearing. We limit our further discussion and analysis to the challenged reliability of the expert opinions the United States expects to adduce from Messrs. Rippman and Wells.

[4] *Daubert v. Merrill Dow Pharmaceuticals, Inc.*, 509 U.S. 579 (1993).

[5] Class characteristics are marks or signs common to an entire line of firearms (e.g. the rifling of all Glock 19 barrels); Subclass Characteristics are marks or signs common to a single run of a particular firearm; Individual Characteristics are marks or signs unique to an individual firearm.

[6] ECF No. 589 at 4, 30.

[7] *See* ECF No. 664.

[8] Exhibit 4.2, AFTE Theory of Identification as it Relates to Toolmarks.

[9] ECF No. 555 at 8-9.

[10] *See generally Daubert*, 509 U.S. 579.

[11] *In re TMI Litigation*, 193 F.3d 613, 663 (3d Cir.1999).

[12] *United States v. Felix*, No. 20-2, 2022 WL 17250458, at *10 (D.V.I. Nov. 28, 2022) (quoting *Pineda v. Ford Motor Co.*, 520 F.3d 237, 247-48 (3d Cir. 2008)).

[13] *See In re Paoli R.R. Yard PCB Litigation (Paoli II)*, 35 F.3d 717, 744 (3d Cir. 1994).

[14] *United States v. Mitchell*, 365 F.3d 215, 244 (3d Cir. 2004) (quoting *Ruiz Troche v. Pepsi Cola of P.R. Bottling Co.*, 161 F.3d 77, 85 (1st Cir. 1998)).

[15] *United States v. Wrensford*, No. 13-3, 2014 WL 3715036, at *17 (D.V.I. July 28, 2014).

[16] *Adams*, 444 F.Supp.3d 1248 (D. Or. 2020).

[17] *Adams*, 444 F.Sup.3d at 1266.

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

[19] *Id.*

[20] *Id.*

[21] *United States v. Pete*, No. 22-48, 2023 WL 4928523 (N.D. Fla. July 21, 2023) (declining to hold a *Daubert* hearing because firearm toolmark analysis "has long been accepted as a field of worthy

expert testimony" and based on parties' submissions permitting expert to testify); *United States v. Randolph*, No. 22-145, 2023 WL 4562855 (M.D. Fla. July 17, 2023) ("A balance of the factors shows that AFTE methodology is sufficiently reliable."); *United States v. Blackman*, No. 18-728, 2023 WL 3440384 (N.D. Ill. May 12, 2023) (holding AFTE methodology of firearms toolmark examination meets reliability factors under *Daubert*); *United States v. Harris,* 502 F.Supp.3d 28 (D.D.C. 2020) (finding firearm and toolmark analysis "reliable and admissible"); *United States v. Johnson*, No. 16-281, 2019 WL 1130258 (S.D.N.Y. Mar. 11, 2019) (balancing *Daubert* factors to find testimony of firearm toolmark examiner admissible); *United States v. Ashburn*, 88 F.Supp.3d 239 (E.D.N.Y. 2015) (finding "firearms identification is at bottom a subjective inquiry" but permitting the expert "to offer an expert opinion regarding a potential ballistics match"); *United States v. Otero*, 849 F.Supp.2d 425 (D.N.J. 2012) (finding expert report and opinion on firearm toolmark analysis "based on a reliable methodology" and admissible under Rule 702); *United States v. Taylor*, 663 F.Supp.2d 1170 (D.N.M. 2009) (expert could testify, in his opinion, using pattern-based methodology, bullet came from suspect rifle to within reasonable degree of certainty in firearms examination field); *United States v. Diaz*, No. 05-167, 2007 WL 485967 (N.D. Cal. Feb. 12, 2007) (holding theory of firearm toolmark identification used by San Francisco Police Crime Lab was reliable under *Daubert*); *United States v. Hicks*, 389 F.3d 514, 526 (5th Cir. 2004) ("[T]he matching of spent shell casings to the weapon that fired them has been a recognized method of ballistics testing in this circuit for decades.").

[22] The United States asserts firearms analysis has been used in administrative proceedings, but also cites *United States v. Wrensford*, No.13-3, 2014 WL 3715036, at *17 (D.V.I. July 28, 2014), where Judge Lewis found such administrative use is not contemplated by this final factor.