

**N THE UNITED STATES DISTRICT COURT  
FOR THE DISTRICT OF KANSAS**

**UNITED STATES OF AMERICA,**

**Plaintiff,**

**v.**

**WARREN RICHARDSON,**

**Defendant.**

**Case No. 19-20076-JAR**

**MEMORANDUM AND ORDER**

This matter comes before the Court on Defendant Warren Richardson's Motion to Exclude or Limit Ballistics Reports and Testimony (Doc. 112). The government has filed a Response in Opposition (Doc. 117). The Court heard testimony from the designated expert, Alexis Lalli, at an evidentiary hearing on November 9, 2023. The Court has considered this evidence, the argument presented by the parties at the hearing, and the parties' briefs, and is now prepared to rule. As explained more fully below, the Court denies the motion to exclude but grants in part the motion to limit reports and testimony.

**I. Background**

Defendant Warren Richardson is charged in a four-count Indictment with (1) conspiring to distribute cocaine base and to possess with intent to distribute cocaine base from about June 1, 2014 to December 17, 2014; (2) unlawfully possessing a firearm on December 13, 2014, after previously being convicted of felony crimes for carrying a firearm in furtherance of drug trafficking and possessing with intent to distribute cocaine base within 1000 feet of a school; (3) possession with intent to distribute cocaine base on December 17, 2014; and (4) using or carrying a firearm on December 17, 2014, in furtherance of drug trafficking.

The two firearms charges in the Indictment are based on Richardson's possession of the same firearm, a black 7.62x39mm assault rifle on (1) December 13, 2014, the day the confidential informant ("informant" or "CI") in this case was murdered, and (2) December 18, when it was seized from the car he was driving at the time of his arrest. The government intends to present evidence that this firearm was used to murder the informant who purchased drugs from Richardson and the coconspirators. The government also intends to present evidence that within two hours after Richardson committed two armed robberies with this firearm, he was arrested after a high-speed chase and officers recovered crack cocaine and the firearm from the vehicle Richardson was driving.

The test-fired cartridge cases from this firearm and the 16 cartridge cases recovered from the December 13 murder scene were submitted for ballistics testing. The ballistics testing, which was conducted in late 2014 or early 2015, determined that tool markings on the cartridge case from a bullet test-fired from the recovered firearm were consistent with the tool markings on 15 of 16 shell casings recovered from the murder scene. The ballistics expert opined that the tool markings on the sixteenth shell casing were an inconclusive match with the tool markings on cartridge casing from the test-fired bullet, but they indicated that the cartridge had been cycled through the recovered firearm.

The government filed a Notice of Expert Witness identifying as a proposed expert, Alexis Lalli, Chief Criminalist Supervisor, Kansas City Missouri Police Department Crime Laboratory-Firearms Section.<sup>1</sup> Lalli testified that since 2014 she has held a certification from the Association of Firearms and Toolmark Examiners ("AFTE") in firearms examination and identification. This certification requires a two-year training program, followed by three years of

---

<sup>1</sup> Doc. 95.

casework performance, written and performance testing, as well as continuing education and periodic recertifications to maintain the certification. The Kansas City, Missouri Department Crime Laboratory is accredited, subject to periodic internal and external audits. Lalli has conducted thousands of ballistics examinations in her career and has testified as an expert in the area of firearms and ballistic analysis in state and federal court many times. In just the last four years she has testified 25 to 30 times.

Lalli testified that Kathy May, another ballistics examiner in her laboratory, first examined the cartridge cases recovered from the murder scene and the cartridge cases from test-fired bullets from the firearm recovered from the vehicle driven by Richardson on December 18. Consistent with the lab's protocol, Lalli then performed a separate, independent examination in her role as a verifier of the examination.

Using a demonstrative exhibit, Lalli described the process of examination. The examiner is primarily looking at the primer area of the cartridge case, which is the part of the cartridge case that comes into contact most with the firearm during the firing process. The examiner uses a comparison microscope, which allows the examiner to look at two cartridges at the same time via a split screen. The examiner then looks at those markings that transfer during the firing process, which are individual characteristics rather than class characteristics, to see if they align and if there is agreement between those, in order to reach an opinion on whether or not the two cartridge cases were fired from the same firearm. Lalli further explained how individual characteristics of a firearm translate to a casing. During the firing process, the cartridge case comes into contact with those areas of the firearm that were manufactured and have individual characteristics, defined as random imperfections that are created during the manufacturing process. The firearm impresses its markings onto the primer portion of the cartridge case. The

individual characteristics of the firearm are unique, for even when firearms are manufactured in an assembly line, the tool used to create the working surfaces of the firearm during the manufacturing process is scraping off metal pieces that break off randomly, leaving markings on the breech face or the firing pin of the firearm. In other words, during the manufacturing process, there are unique markings left on the breech face and/or firing pin because the manufacturing tool, each time it is used, changes just slightly as it scrapes off metal from one breech face to another or one firing pin to another, and thus creates random imperfections with the manufacture of each firearm.

Lalli testified that she started with the 16 cartridge cases and first looked at the class characteristics because if there is not a match of the class characteristics, the analysis goes no further. Class characteristics are determined prior to the manufacturing process and include such characteristics as the caliber and shape of the firing pin.

Next, Lalli proceeded to examine the individual characteristics of the cartridge cases. Lalli explained that individual characteristics are marks produced by the random imperfections of tool surfaces created during the manufacturing process, and/or by use, corrosion, or damage to the firearm. Individual characteristics, such as striations, patterns, and/or groupings of marks are unique to that firearm and distinguish it from all other firearms. Lalli testified that she compared all sixteen cartridge cases recovered from the murder scene and that fifteen of these cartridge cases were fired from the same firearm. The examination is done by using a comparison microscope that has a split screen allowing the examiner to view the toolmarkings on two compared-cartridge cases side-by-side. Lalli testified that this is a visual and subjective examination that focuses on whether the toolmarkings on the compared subjects are aligned and in agreement, which she characterized as objective criteria.

Lalli's examination determined that for 15 of the 16 cartridge cases recovered from the murder scene, there was agreement of all discernible class characteristics and sufficient agreement of a combination of individual characteristics, including breech face marks<sup>2</sup> and aperture shear<sup>3</sup> for her to conclude that the toolmarks on these 15 cartridge cases were fired from the same firearm. By sufficient agreement, she means that the likelihood another firearm could have made the tool mark is so remote as to be considered a practical impossibility.

With respect to the sixteenth cartridge, Lalli testified that it was inconclusive that it was fired from same weapon as the other 15 cartridges because there was an insufficient quantity and quality of individual characteristics present. When an examiner has an inconclusive result, they then examine to see if they can find any cycle (chambered, extracted or ejected) of firearms in common between the cartridge cases. Lalli determined that on the sixteenth cartridge case, there was an extractor mark that had agreement with extractor marks on the other 15 cartridge cases leading her to conclude that the sixteenth cartridge case was cycled through the same firearm as the other 15, based on areas of agreement found in the extractor marks.<sup>4</sup>

Having opined that 15 of the cartridge cases recovered from the murder scene were fired from the same firearm, Lalli then examined the individual characteristics on the cartridge cases from four bullets test-fired from the recovered firearm. She then compared the individual characteristics present on one of the test-fired cartridge cases to the individual characteristics on

---

<sup>2</sup> Lalli explained that breech face is the rear portion of the firearm or bullet that rests on the head of the cartridge once it is in the chamber of the firearm. When the bullet is fired, the cartridge case slams back against the breech or bolt of the firearm and picks up markings from that breech or bolt. Then the cartridge case is extracted and ejected out of the firearm.

<sup>3</sup> Lalli explained that aperture shear is the area that occurs between the firing pin hole and the breech. When the bullet is fired, some of the cartridge case primer can melt back into the firing pin hole, and then during the extraction process, it can create a shearing motion that will cause shear marks or striated marks on the cartridge case primer.

<sup>4</sup> Lalli explained that when a cartridge case from an unfired or fired bullet is cycled through a firearm, there may be markings on the cartridge case from the extraction and/or ejection process.

the 15 cartridge cases. Based on this comparison, Lalli determined that the 15 cartridge cases were fired from the same firearm as the test-fired cartridge cases, and the sixteenth cartridge case was cycled through the same firearm as the test-fired cartridges.

Thus, in Lalli's opinion, the 15 cartridge cases recovered from the murder scene were fired from the same firearm recovered from the stolen vehicle Richardson was driving on December 18, five days after the murder. Lalli testified that, based on the agreement of all discernible class characteristics and sufficient agreement of a combination of individual characteristics, including the breech face marks and aperture shear, the extent of the agreement exceeds that which can occur in the comparison of toolmarks made by different firearms. She further opined that with respect to the sixteenth cartridge case recovered from the murder scene, areas of agreement were found in the extractor marks such that that cartridge case was cycled through the same firearm recovered from the vehicle driven by Richardson.

On cross-examination, Lalli acknowledged that a 2009 report by the National Academy of Sciences<sup>5</sup> and a 2016 report from the President's Council of Advisors on Science and Technology ("PCAST Report")<sup>6</sup> were critical of the scientific criteria for foundational validity and the scientific underpinnings of toolmark discipline. Lalli testified that one criticism—that there is a lack of statistical data—has been addressed by ongoing studies since these reports that demonstrate that there are low error rates of less than one to two percent, such that the ballistic examinations are reliable.

---

<sup>5</sup> National Research Council, *Strengthening Forensic Science in the United States: A Path Forward* (2009).

<sup>6</sup> President's Council of Advisors on Sci. & Tech., Exec. Off. of the President, *Report to the President: Forensic Science in Criminal Courts: Ensuring Scientific Validity of Feature-Comparison Methods* 63 (Sept. 2016), [https://obamawhitehouse.archives.gov/sites/default/files/microsites/ostp/PCAST/pcast\\_forensic\\_science\\_report\\_final.pdf](https://obamawhitehouse.archives.gov/sites/default/files/microsites/ostp/PCAST/pcast_forensic_science_report_final.pdf) [hereinafter PCAST Report].

## II. Standard

Richardson moves to exclude the ballistics reports and Lalli's testimony under Fed. R. Evid. 702 and *Daubert v. Merrell Dow Pharmaceuticals, Inc.*<sup>7</sup> Because the government responds that it will not be introducing the ballistics reports into evidence at trial, the Court confines its ruling to the motion to exclude Lalli's testimony.

Rule 702, which was recently amended effective December 1, 2023, governs the admissibility of expert witness testimony by allowing someone "who is qualified as an expert by knowledge, skill, experience, training, or education [to] testify in the form of an opinion."<sup>8</sup> The proponent of the expert's opinion must

demonstrate[] to the court that it is more likely than not that:

- (a) the expert's scientific, technical, or other specialized knowledge will help the trier of fact to understand the evidence or to determine a fact in issue;
- (b) the testimony is based on sufficient facts or data;
- (c) the testimony is the product of reliable principles and methods; and
- (d) the expert's opinion reflects a reliable application of the principles and methods to the facts of the case.<sup>9</sup>

In *Daubert*, the Supreme Court held that Rule 702 imposes a gatekeeping responsibility on trial courts to ensure that proposed expert testimony "is not only relevant, but reliable."<sup>10</sup> In performing this gatekeeping function, the court "generally must first determine whether the expert is qualified 'by knowledge, skill, experience, training, or education' to render an

---

<sup>7</sup> 509 U.S. 579 (1993).

<sup>8</sup> Fed. R. Evid. 702.

<sup>9</sup> *Id.*

<sup>10</sup> 509 U.S. at 589.

opinion.”<sup>11</sup> If the expert is sufficiently qualified, the court must next determine whether the expert’s testimony has “a reliable basis in the knowledge and experience of [the relevant] discipline.”<sup>12</sup> The Supreme Court in *Daubert* set forth a non-exhaustive list of four factors that courts may consider in determining the reliability of the proffered expert testimony: (1) whether the theory or technique can be and has been tested; (2) whether it has been subjected to peer review and publication; (3) the known or potential error rate; and (4) its degree of general acceptance in the relevant scientific community.<sup>13</sup>

After determining that a witness is qualified to testify as an expert and that the testimony is reliable, the court must determine whether the expert testimony is sufficiently “relevant to the task at hand.”<sup>14</sup> Under Fed. R. Evid. 401, evidence is relevant if “it has any tendency to make a fact more or less probable than it would be without the evidence,” and “the fact is of consequence in determining the action.”<sup>15</sup> “Relevant expert testimony must ‘logically advance[] a material aspect of the case’ and be ‘sufficiently tied to the facts of the case that it will aid the jury in resolving a factual dispute.’”<sup>16</sup> “Expert testimony which does not relate to any issue in the case is not relevant and, ergo, non-helpful.”<sup>17</sup> In assessing whether expert testimony will

---

<sup>11</sup> *United States v. Nacchio*, 555 F.3d 1234, 1241 (10th Cir. 2009) (en banc) (quoting Fed. R. Evid. 702).

<sup>12</sup> *Bitler v. A.O. Smith Corp.*, 400 F.3d 1227, 1232 (10th Cir. 2005) (quoting *Daubert*, 509 U.S. at 592).

<sup>13</sup> *Daubert*, 509 U.S. at 593–94.

<sup>14</sup> *Bitler v. A.O. Smith Corp.*, 400 F.3d 1227, 1234 (10th Cir. 2005) (quoting *Daubert*, 509 U.S. at 597).

<sup>15</sup> Fed. R. Evid. 401; *see* Fed. R. Evid. 402 (“Irrelevant evidence is not admissible.”).

<sup>16</sup> *United States v. Garcia*, 635 F.3d 472, 476 (10th Cir. 2011) (alteration in original) (first quoting *Norris v. Baxter Healthcare Corp.*, 397 F.3d 878, 884 n.2 (10th Cir. 2005); and then quoting *Daubert*, 509 U.S. at 591).

<sup>17</sup> *Daubert*, 509 U.S. at 591 (quoting 3 Weinstein & Berger ¶ 702[02] (1988)).



assist the jury, the court should consider whether the testimony “is within the juror’s common knowledge and experience.”<sup>18</sup>

It is within the court’s discretion to determine how to perform its gatekeeping function under *Daubert*.<sup>19</sup> The most common method for fulfilling this function is a *Daubert* hearing, although it is not specifically mandated.<sup>20</sup> While the November 9, 2023 hearing was not noticed as a *Daubert* hearing, Lalli testified and Richardson cross-examined her, challenging her testimony and the discipline as unreliable, untested, subject to high error rates, not peer reviewed, and roundly criticized by the PCAST and National Academy of Sciences reports. The November 9, 2023 hearing on the motions in limine was effectively a *Daubert* hearing because it constituted an evidentiary hearing on the instant motion. The Court granted the parties’ requests to supplement the record after the hearing and neither party requested further hearing. The record is thus now closed for purposes of this motion.

### **III. Discussion**

#### **A. Qualifications**

Defendant Richardson does not question Ms. Lalli’s qualifications, though he reserves his right to object at trial to her qualifications under Rule 702. Having received Lalli’s curriculum vitae,<sup>21</sup> and having heard her testimony during the November 9, 2023 hearing, the Court finds that she is qualified as an expert under Rule 702.

---

<sup>18</sup> *Garcia*, 635 F.3d at 476–77 (quoting *United States v. Rodriguez-Felix*, 450 F.3d 1117, 1123 (10th Cir. 2006)).

<sup>19</sup> *Goebel v. Denver & Rio Grande W. R.R.*, 215 F.3d 1083, 1087 (10th Cir. 2000) (citing *Kumho Tire Co. v. Carmichael*, 526 U.S. 137, 152 (1999)).

<sup>20</sup> *Id.*

<sup>21</sup> Doc. 95-1.

## **B. Reliability**

In determining the reliability of proffered expert testimony, the Supreme Court has emphasized that the four reliability factors recited above are not a “definitive checklist or test” and that a court’s gatekeeping inquiry into reliability must be “tied to the facts of a particular case.”<sup>22</sup> In some cases, “the relevant reliability concerns may focus upon personal knowledge or experience,” rather than the *Daubert* factors.<sup>23</sup> Quite simply, under Rule 702, the reliability criterion remains a discrete, independent, and important requirement for admissibility.”<sup>24</sup> Ultimately, the court’s role “is to make certain that an expert, whether basing testimony upon professional studies or personal experience, employs in the courtroom the same level of intellectual rigor that characterizes the practice of an expert in the relevant field.”<sup>25</sup>

### **1. Testability**

The first reliability factor under *Daubert* is whether the theory or technique can be and has been tested.<sup>26</sup> Defendant argues that the theory or technique is not testable or tested. As described by Lalli, and as described in many of the sources cited by the parties, including reports, studies, and court decisions,<sup>27</sup> the AFTE’s position is that a qualified firearms examiner can determine whether multiple cartridges were fired from the same firearm by comparing toolmarks on the cartridges. The AFTE technique compares class characteristics to determine if they are the same, and if so, then uses a comparison microscope to compare individual characteristics to look for alignment and agreement of toolmark striae. The AFTE standard is

---

<sup>22</sup> *Kumho Tire Co. v. Carmichael*, 526 U.S. 137, 150 (1999).

<sup>23</sup> *Id.*

<sup>24</sup> *United States v. Frazier*, 387 F.3d 1244, 1261 (11th Cir. 2004) (emphasis omitted).

<sup>25</sup> *Kumho Tire Co.*, 526 U.S. at 152.

<sup>26</sup> *Daubert*, 509 U.S. at 593–94.

<sup>27</sup> *See United States v. Hunt*, 63 F.4th 1229, 1235–36 (10th Cir. 2023).

that there must be “[a]greement of all discernible class characteristics and sufficient agreement of a combination of individual characteristics where the extent of agreement exceeds that which can occur in the comparison of toolmarks made by different tools and is consistent with the agreement demonstrated by toolmarks known to have been produced by the same tool.”<sup>28</sup>

As the Tenth Circuit recently noted in *United States v. Hunt*,<sup>29</sup> the comparison of toolmarks on ammunition to determine whether it was expended from the same firearm has been used for over a century, and advances in science and technology have refined the field.<sup>30</sup> But, beginning about 15 years ago, criticism by well-respected bodies has subjected the field to closer scrutiny.<sup>31</sup> In 2009, the National Academy of Sciences issued a report (“NRC Report”)<sup>32</sup> that recommended further research be done in multiple forensic science disciplines, including firearm toolmark analysis, to address issues concerning accuracy, reliability, and validity.<sup>33</sup>

The NRC Report did not conclude that firearm toolmark analysis was wholly invalid; it found that class characteristics were helpful in identifying the pool of tools that may have left a distinctive mark and that individual characteristics might be distinctive enough to suggest one particular source.<sup>34</sup> But the NRC Report found that there had not been sufficient studies of the reliability and repeatability of AFTE’s methods and that AFTE lacked a precisely defined process and specific protocol.<sup>35</sup>

---

<sup>28</sup> *Id.* at 1236 (quoting the Ass’n of Firearm & Tool Mark Exam’rs, *Glossary* 94 (6th ed. 2013)).

<sup>29</sup> *Id.* at 1235 (citations omitted).

<sup>30</sup> *Id.*

<sup>31</sup> *Id.* at 1236.

<sup>32</sup> National Research Council, *Strengthening Forensic Science in the United States: A Path Forward* (2009).

<sup>33</sup> *Id.* at 1–2.

<sup>34</sup> *Id.* at 154.

<sup>35</sup> *Id.* at 155.

In 2016, the President’s Council of Advisors on Science and Technology issued a report critical of the scientific criteria for foundational validity and the scientific underpinnings of firearm toolmark examination, finding that it has “lacked scientific, statistical proof that would independently corroborate conclusions based on morphology characteristics.”<sup>36</sup> The PCAST Report also called for additional studies to verify the principles and methods underlying toolmark examination.<sup>37</sup>

The government points the Court to AFTE’s website, which maintains a nonexhaustive list of what it considers to be the more important studies on “Testability of the Scientific Principle” underlying AFTE’s theory, methodology and standards.<sup>38</sup> Most of these studies are cited to in the Journal of Forensic Science or the AFTE Journal. There are numerous such studies, most of which predate the PCAST Report, including those listed on AFTE’s website: 28 studies in the category of Firearm Identification—Cartridge Cases, including the 2014 Ames Laboratory Study; 9 studies in the category of Firearm and Toolmark Identification—Theoretical; 30 studies in the category of Toolmark Identification, including one study since the PCAST Report; and 6 studies under the category of Emerging Research.<sup>39</sup>

The PCAST Report criticized the large body of prior validation studies as “inappropriately designed to assess foundational validity and estimate reliability.”<sup>40</sup> It opined that there had only been one study that was an appropriately designed “black-box” study of

---

<sup>36</sup> See PCAST Report, *supra* note 6.

<sup>37</sup> *Id.* at 11–12.

<sup>38</sup> AFTE, Resources, <https://afte.org/resources/swggun-ark/testability-of-the-scientific-principle> (last visited Feb. 9, 2024).

<sup>39</sup> *Id.*

<sup>40</sup> PCAST Report, *supra* note 6, at 11.

firearms: the Ames Laboratory Study, which was conducted by an independent testing lab.<sup>41</sup>

This Court declines to decide whether the PCAST-dictated design for such studies is exclusively correct, and declines to decide whether Rule 702 or *Daubert* requires that particular design of a study in this field. Instead, the Court focuses on subsequent studies that do conform to the black-box design urged in the PCAST report.

The government cites to *United States v. Harris*,<sup>42</sup> in which the court discussed these three post-PCAST black-box studies of the validity and reliability of firearm toolmark examination: (1) the 2018 Heat Map Study, (2) the Lilien Study, and (3) the Keisler Study. All three studies evidenced that firearm and toolmark identification can be tested, and further can be reasonably assessed for reliability. And, as addressed below, these studies, along with the large body of other studies that the PCAST report dismissed, strongly suggest reliability because of low, single-digit error rates. The government further cites to a 2022 law review article by James Agar, an assistant general counsel for the FBI's forensic laboratory at Quantico, that discusses two more post-PCAST "black-box" studies.<sup>43</sup>

Richardson poses two related arguments under the testability factor: that firearm toolmark examination is subjective and cannot be tested and that the AFTE is not based on a valid standard. But as Lalli explained, her analysis is based on objective criteria—alignment and agreement of markings or striae on the cartridges under comparison. While the agreement in such markings or striae may be a subjective call, it is one based on extensive training and experience of the examiner. The mere fact that there is a subjective component, similar to

---

<sup>41</sup> *Id.*

<sup>42</sup> 502 F. Supp. 3d 28, 37–38 (D.D.C. 2020).

<sup>43</sup> See James Agar, *The Admissibility of Firearms and Toolmarks Expert Testimony in the Shadow of PCAST*, 74 Baylor L. Rev. 93, 169–71 (2022).

fingerprint analysis, does not mean it is not testable; indeed, it has been repeatedly tested. Further, the fact that the examiners must document their results and findings through written report and photographs and that, at least in Lalli's accredited lab, there have to be two examinations done by independent examiners, negates the concern that the examination is merely a subjective guess or speculation. Moreover, Richardson can cross-examine the expert about any inherent subjectivity in firearm toolmark examination.<sup>44</sup>

Richardson further argues that toolmark identification lacks reliable standards controlling the technique's operation and that Lalli's AFTE certification does not suffice, given that the PCAST report found that AFTE's stated method and standard is circular.<sup>45</sup> The criticism that the method and standard is circular is based on AFTE's declaration that an examiner may state that two toolmarks have a common origin when their features are in "sufficient agreement," and then defines "sufficient agreement" as occurring when the examiner considers it a "practical impossibility" that the toolmarks have different origins."<sup>46</sup>

To be sure, firearm toolmark examination does not lend itself to a numerically stated result. But that does not mean the discipline is not governed by standards. *Daubert* counsels the gatekeeping court to focus on standards "governing the technique's operation,"<sup>47</sup> not on whether there are numerical or statistical results.<sup>48</sup> The government argues that the AFTE method has

---

<sup>44</sup> *United States v. Gil*, 680 F. App'x 11, 13 (2d Cir. 2017) ("[A]rguments about the subjectivity inherent in otherwise reliable methodologies go 'to the weight of the evidence, not to its admissibility,' and [are] 'matters for cross-examination and argument to the jury'" (quoting *United States v. Romano*, 794 F.3d 317, 333 (2d Cir. 2015))).

<sup>45</sup> See PCAST Report, *supra* note 6, at 60.

<sup>46</sup> *Id.*

<sup>47</sup> *Daubert v. Merrell Dow Pharms., Inc.*, 509 U.S. 597, 594 (1993).

<sup>48</sup> *United States v. Hunt*, 464 F. Supp. 3d 1252, 1260 (W.D. Okla. 2020) ("*Daubert* does not impose a rigid requirement that the expert reach a conclusion through an entirely objective set of criteria."), *aff'd*, 63 F.4th 1229, 1235 (10th Cir. 2023); *United States v. Harris*, 502 F. Supp. 3d 28, 42 (D.D.C. 2020) ("Rule 702 'does not impose a requirement that the expert must reach a conclusion via an objective set of criteria or that he be able to quantify his opinion with a statistical probability.'").

industry standards, including the laboratory’s standard operating procedures and guidelines as well as training, proficiency testing, and validation procedures.<sup>49</sup> The Court should not conflate standards governing operation of a technique with a desire for an objective or numerically defined result.

Given this history, the ongoing developments in the field, and the record in this case, this Court is persuaded that firearm toolmark examination can be and has been sufficiently tested for validity and reliability to meet the requirements of Rule 702 and *Daubert*. Neither the NRC Report nor the PCAST Report called for the complete exclusion of such evidence and this Court has found no other court that has excluded such evidence on the basis that the discipline cannot or has not been tested for validity and reliability.<sup>50</sup>

Yet, as the Tenth Circuit cautioned in *Hunt*, even while affirming the district court’s admission of the evidence under Rule 702 and *Daubert*, in light of the critiques expressed in the PCAST and NRC reports, courts should be mindful of not ruling too broadly on this issue.<sup>51</sup> As addressed below, the Court limits the language that the expert may use in opining on her findings.

---

<sup>49</sup> See *United States v. Rhodes*, No. 19-0333-MC, 2023 WL 196174, at \*5 (D. Or. Jan. 17, 2023) (citing the “industry standards that safeguard the process,” which “include: a specific laboratory’s standard operating procedures and guidelines; International Organization for Standardization (“ISO”)/International Electrotechnical Commission (“IEC”) Standard 17025; training, monitoring, validation of procedures, and regular proficiency testing.”).

<sup>50</sup> See, e.g., *United States v. Brown*, 973 F.3d 667, 702–04 (7th Cir. 2020) (finding that the district court did not abuse its discretion in determining that toolmark evidence admissible because the AFTE method had been tested, peer-reviewed, had overall low error rate and is widely-accepted in the judicial system and beyond); *United States v. Johnson* 875 F.3d 1265, 1281 (9th Cir. 2017) (finding no abuse of discretion in denying *Daubert* motion relying on a number of cases and scientific sources establishing that the AFTE methodology satisfies *Daubert* and no defendant could cite to no case that had ever excluded AFTE ballistics testimony altogether).

<sup>51</sup> 63 F.4th 1229, 1244–45 (10th Cir. 2023).

## 2. Peer Review and Publication

The second reliability factor is whether the AFTE theory or technique has been subjected to peer review and publication. As detailed above, the AFTE website collects some of the literature citations to studies on firearm identification. Many of the citations are to articles published in the Journal of Forensic Science or the AFTE Journal, both of which are peer-reviewed.<sup>52</sup> At least three other peer-reviewed journals address the AFTE method.<sup>53</sup> “[N]umerous courts have concluded that publication in the AFTE Journal satisfies this prong of the *Daubert* admissibility analysis.”<sup>54</sup> The Court easily concludes that this *Daubert* reliability factor has been met.

## 3. Known or Potential Error Rate

The third factor is whether the AFTE theory or technique has a known or potential error rate. After the NRC Report in 2009, the 2014 Ames Study established known or potential error rates.<sup>55</sup> While the 2016 PCAST Report noted that the 2014 Ames Study had not been published or peer-reviewed,<sup>56</sup> the PCAST Report nonetheless viewed the 2014 Ames Study favorably as

---

<sup>52</sup> See *United States v. Otero*, 849 F. Supp. 2d 425, 433 (D.N.J. 2012) (explaining that the AFTE theory is subject to peer review through submission of validation studies to and publication by the AFTE Journal, which has a formal process for writing and submitting manuscripts, assignment of manuscripts to other experts within the scientific community for a technical review, returning of manuscripts to authors for clarification or re-write, and a final review by the Editorial Committee, as well as a formal post-publication peer review process, allowing AFTE members and any other interested individuals to comment on previously published articles).

<sup>53</sup> See *Brown*, 973 F.3d at 704 (citing three different peer-reviewed journals that address the AFTE method).

<sup>54</sup> *Harris*, 502 F. Supp. 3d at 40 (citations omitted).

<sup>55</sup> Daniel P. Baldwin et al., *A study of False-Positive and False-Negative Error Rates in Cartridge Case Comparisons*. Ames Lab’y, USDOE Technical Report #IS-5207 (Apr. 7, 2014).

<sup>56</sup> As Richardson points out, another study by the Ames Laboratory in 2020 evaluated accuracy of firearms examinations. This study was peer-reviewed, and one peer reviewer was critical that the 2020 Ames Study had weak repeatability and reproducibility. But the study itself did not indicate high error rates. See Alan H. Dorfman & Richard Valliant, *A Re-Analysis of Repeatability and Reproducibility in the Ames-USDOE-FBI Study*, 9 Statistics & Pub. Policy, no. 1, 2022, at 175–84, <https://www.tandfonline.com/doi/epdf/10.1080/2330443X.2022.2120137?needAccess=true>; Keith L. Monson et al., *Planning, Design and Logistics of a Decision Analysis Study: The*



the sole black-box designed study at that time. In the 2014 Ames Study, 218 examiners were presented with 15 separate comparison problems.<sup>57</sup> The study concluded that there were low error rates in firearm toolmark analysis, specifically a 1.5% false positive rate in conclusive examinations.<sup>58</sup>

The PCAST Report of course recommended that further black-box studies be done. That has since occurred and post-PCAST studies continue to evidence low error rates. The three post-PCAST studies discussed by the court in *United States v. Harris*,<sup>59</sup> all had very low rates of error. Two studies showed an error rate of 0% and the third a false positive rate of 0.43%.<sup>60</sup> In the two additional post-PCAST studies discussed in Agar's article, one study showed an overall error rate of 0.16% and the other a false positive rate of 0.656% for bullets, and 0.933% for cartridge cases.<sup>61</sup> The government points the Court to another recent post-PCAST black-box study on the accuracy of conclusions by firearm and toolmark examiners, which estimated that the overall rate of false identifications is 1.01%.<sup>62</sup>

Moreover, this Court agrees with Lalli, Agar,<sup>63</sup> as well as other courts<sup>64</sup> that note these studies may overstate the error rate. For these studies do not account for the quality assurance and control processes used by laboratories to mitigate any potential error by individual

---

*FBI/Ames Study Involving Forensic Firearms Examiners*, 4 Forensic Sci. Int'l 4 (2022) (explaining that the recent FBI/Ames study produced an estimated false positive error rate in the 0.933–1.57% range).

<sup>57</sup> PCAST Report, *supra* note 6, at 110.

<sup>58</sup> *Id.*; see also Monson, *supra* note 56, at 4 (collecting studies that place the error rate in low single digits).

<sup>59</sup> *Harris*, 502 F. Supp. 3d at 39.

<sup>60</sup> *Id.*

<sup>61</sup> See Agar, *supra* note 43, at 170–71.

<sup>62</sup> David P. Baldwin et al., *A Study of Examiner Accuracy in Cartridge Case Comparisons Part 1: Examiner Error Rates*, 349 Forensic Sci. Int'l (2023).

<sup>63</sup> See Agar, *supra* note 43, at 168–71.

<sup>64</sup> See *United States v. Rhodes*, No. 19-0333-MC, 2023 WL 196174, at \*4 (D. Or. Jan. 17, 2023)

examiners, such as the process used by Lalli’s AFTE accredited lab of having the evidence submitted to a second, independent examiner for verification, such as there was in this case. Further, as Agar noted, the error rate may well be overstated given that other design elements in these studies do not replicate what happens in the field.<sup>65</sup> Many of these studies also “incorporate the examination of bullets and cartridge cases fired from consecutively manufactured barrels or slides where sub-class characteristics are present and which could potentially mislead the examiner.”<sup>66</sup> “This particular design factor in these studies makes firearms identification studies more difficult than the vast majority of casework from real-life shootings, where criminals do not use multiple firearms equipped with consecutively made barrels or slides, nor the same caliber of firearms or even firearms from the same manufacturer.”<sup>67</sup>

#### **4. Degree of General Acceptance in the Relevant Scientific Community**

The final reliability factor under *Daubert*’s nonexclusive list of factors is the degree of general acceptance in the relevant scientific community. This factor is easily satisfied. The relevant scientific field is firearms examiners. The field has appropriately faced criticism and closer scrutiny. But the field, as well as the academy, has and continues to appropriately respond with more studies and more stringently designed studies of this important area of forensic science. The field still enjoys general acceptance among firearms examiners.<sup>68</sup>

---

<sup>65</sup> Agar, *supra* note 43, at 140.

<sup>66</sup> *Id.* at 168.

<sup>67</sup> *Id.* at 168–69.

<sup>68</sup> *United States v. Hunt*, 464 F. Supp. 3d 1252, 1259 (W.D. Okla. 2020) (citing *United States v. Romero-Lobato*, 379 F. Supp. 3d 1111 1121 (D. Nev. 2019)); *Romero-Lobato*, 379 F. Supp. 3d 1121 (collecting cases).

In sum, the Court finds that Lalli’s testimony satisfies the reliability components of Rule 702 and *Daubert*. For the reasons stated above, it contains sufficient indicia of reliability to be admissible at trial, subject to the limitations discussed below.

**C. Relevance**

Richardson does not contest that Lalli’s testimony would assist the jury in determining a fact in issue at trial, and the Court easily finds that whether Richardson unlawfully possessed the firearm in question on December 13, the date of the informant’s murder, is relevant to the charges in this case.

**D. Limitations of Expert’s testimony**

Richardson argues in the alternative that if Lalli is permitted to testify, limitations should be imposed on her testimony. He argues that Lalli should be precluded from “expressing any level of certainty about her opinion”<sup>69</sup> and that her testimony should be “limited to opining that the recovered casings were consistent with having been fired from the model of firearm recovered three days later.”<sup>70</sup>

The government represents that it will not elicit testimony on direct examination that the cartridge casings in question match the test-fired casings with absolute or 100% certainty, or even to a reasonable degree of scientific certainty; nor that the firearm and toolmark identification discipline is infallible.<sup>71</sup> Indeed, the Department of Justice (“DOJ”) has issued guidelines to its prosecutors and witnesses, called United States Department of Justice Uniform Language for Testimony and Reports for the Forensic Firearms/Toolmarks Discipline–Pattern

---

<sup>69</sup> Doc. 112 at 7.

<sup>70</sup> *Id.* at 9.

<sup>71</sup> Doc. 117 at 14.

Examination (“DOJ Guidelines”).<sup>72</sup> Effective August 16, 2023, the DOJ Guidelines state in pertinent part:

**Source Identification**

‘Source identification’ is an examiner’s conclusion that two toolmarks originated from the same source. This conclusion is an examiner’s opinion that all observed class characteristics are in agreement and the quality and quantity of corresponding individual characteristics is such that the examiner would not expect to find that same combination of individual characteristics repeated in another source and has found insufficient disagreement of individual characteristics to conclude they originated from different sources.

The basis for a ‘source identification’ conclusion is an examiner’s opinion that the observed class characteristics and corresponding individual characteristics provide extremely strong support for the proposition that the two toolmarks originated from the same source and extremely weak support for the proposition that the two toolmarks originated from different sources.

A ‘source identification’ is the statement of an examiner’s opinion (an inductive inference) that the probability that the two toolmarks were made by different sources is so small that it is negligible.<sup>73</sup>

The DOJ Guidelines provide further:

**IV. Qualifications and Limitations of Forensic Firearms/Toolmarks Discipline Examinations**

- A conclusion provided during testimony or in a report is ultimately an examiner’s decision and is not based on a statistically-derived or verified measurement or comparison to all other firearms or toolmarks. Therefore, an examiner shall not:
  - assert that a ‘source identification’ or a ‘source exclusion’ conclusion is based on the ‘uniqueness’ of an item of evidence.

---

<sup>72</sup> U.S. Dep’t of Justice, *Uniform Language for Testimony and Reports for the Forensic Firearms/Toolmarks Discipline – Pattern Match Examination* (effective Aug. 16, 2023), <https://www.justice.gov/olp/uniform-language-testimony-and-reports>.

<sup>73</sup> *Id.* at 2 (footnote omitted).

- use the terms ‘individualize’ or ‘individualization’ when describing a source conclusion.
  - assert that two toolmarks originated from the same source to the exclusion of all other sources.
- An examiner shall not assert that examinations conducted in the forensic firearms/toolmarks discipline are infallible or have a zero error rate.
  - An examiner shall not provide a conclusion that includes a statistic or numerical degree of probability except when based on relevant and appropriate data.
  - An examiner shall not cite the number of examinations conducted in the forensic firearms/toolmarks discipline performed in his or her career as a direct measure for the accuracy of a conclusion provided. An examiner may cite the number of examinations conducted in the forensic firearms/toolmarks discipline performed in his or her career for the purpose of establishing, defending, or describing his or her qualifications or experience.
  - An examiner shall not assert that two toolmarks originated from the same source with absolute or 100% certainty, or use the expressions ‘reasonable degree of scientific certainty,’ ‘reasonable scientific certainty,’ or similar assertions of reasonable certainty in either reports or testimony unless required to do so by a judge or applicable law.<sup>74</sup>

This Court will follow the strictures in the Department of Justice guidelines, finding that the language it authorizes and the language it precludes is reasonable, given the state of the discipline of firearm toolmark examination. To the extent Defendant Richardson believes Ms. Lalli’s opinion is unsound, he is free to challenge it through cross-examination, as he did at the evidentiary hearing, or by presenting testimony from any expert he has designated.

---

<sup>74</sup> *Id.* at 3 (footnote omitted).

#### **IV. Conclusion**

The Court denies the motion to exclude the testimony of the government's firearms examiner Alexis Lalli, finding that the government has met its burden of proving by a preponderance of the evidence that such testimony is admissible under Rule 702 and *Daubert*. However, the Court places limitations on the substance of Lalli's testimony, adopting in total the currently effective strictures in the DOJ Guidelines for this type of testimony.

**IT IS THEREFORE ORDERED BY THE COURT** that Defendant's Motion to Exclude or Limit Ballistics Reports and Testimony (Doc. 112) is **granted in part and denied in part** as explained in this Order.

**IT IS SO ORDERED.**

Dated: March 6, 2024

S/ Julie A. Robinson  
JULIE A. ROBINSON  
UNITED STATES DISTRICT JUDGE