

FILED
United States Court of Appeals
Tenth Circuit

PUBLISH

March 24, 2023

UNITED STATES COURT OF APPEALS

Christopher M. Wolpert
Clerk of Court

FOR THE TENTH CIRCUIT

UNITED STATES OF AMERICA,

Plaintiff - Appellee,

v.

No. 21-6046

DOMINIC EUGENE HUNT, a/k/a Dime
Sack,

Defendant - Appellant.

**Appeal from the United States District Court
for the Western District of Oklahoma
(D.C. No. 5:19-CR-00073-R-1)**

Grant R. Smith, Assistant Federal Public Defender (Virginia L. Grady, Federal Public Defender, with him on the briefs), Denver, Colorado, for Defendant-Appellant.

Jacquelyn M. Hutzell, Assistant United States Attorney (Robert J. Troester, United States Attorney, and David McCrary, Assistant United States Attorney, with her on the brief), Oklahoma City, Oklahoma, for Plaintiff-Appellee.

Before **HARTZ, SEYMOUR**, and **MORITZ**, Circuit Judges.

HARTZ, Circuit Judge.

Defendant Dominic Eugene Hunt appeals his convictions on two charges of being a felon in possession of ammunition. The ammunition was used in two shootings in early 2019. Investigators found three spent cartridges at the scene of one shooting and one

spent cartridge at the other. A firearms expert testified that all four cartridges were fired from the same (undiscovered) weapon. Defendant's sole complaint on appeal is that the expert testimony should not have been admitted at trial. He argues that the expert's field of firearm toolmark examination is not scientifically valid and that the district court failed to perform its gatekeeping role in examining the admissibility of expert testimony because it relied on prior judicial opinions rather than the most up-to-date empirical evidence when it denied his pretrial motion to exclude the testimony without conducting a hearing.

Exercising jurisdiction under 28 U.S.C. § 1291, we affirm. We need not declare a general rule on the admission of firearm toolmark testimony. We hold only that the district court adequately performed its gatekeeping role and did not err in admitting the testimony in light of the material presented on the pretrial motion and the expert testimony at trial.

I. BACKGROUND

On January 20, 2019, a car was stolen from the residence of Defendant's cousin, Jimmy Jones. The theft was captured on the surveillance camera at Jones's home. It showed that after his daughter started her car to warm it up and went back inside, a man exited a blue Hyundai that had driven by and then jumped in her car and drove off following the Hyundai. Jones later thought he saw the same blue Hyundai down the street, though it turned out that he was mistaken. Based on this misidentification, however, he, Defendant, Travis Carter (Jones's brother), and Christopher Dawson (Defendant's brother) confronted three men that Jones thought

were involved in the theft. This confrontation led to a fistfight, which ended when someone shot Del Lavar Brison, one of the men from the group that Defendant confronted. An officer with the Oklahoma City Police Department (OCPD) was called to the scene and asked Brison who shot him. Brison said something that the responding officer understood as indicating that “it was a black male wearing a . . . maroon jacket.” R., Vol. III at 213. The surveillance video from Jones’s home showed Defendant wearing a maroon hoodie at the time of the theft. An OCPD crime-scene investigator recovered one spent Blazer 9mm Luger cartridge case from the scene—it was the only cartridge case that was found.

Less than two weeks later, in the early morning of February 2, 2019, a man named Conilius Wright was found unconscious in his truck after being mortally wounded in a drive-by shooting. One of Wright’s companions testified that Defendant and Wright had issues with one another, and Defendant’s then girlfriend testified that on the night of the shooting Wright spoke with her about possibly being the father of her child. Defendant’s cell-phone location data indicated that he was near Wright’s shooting one minute before it was reported on a 911 call. An OCPD crime-scene investigator recovered three spent cartridge cases near Wright’s vehicle: one Blazer 9mm Luger cartridge case and two Winchester 9mm Luger cartridge cases.

The four 9mm Luger cartridge cases recovered from the January and February 2019 incidents were submitted to the OCPD Firearms Laboratory and were analyzed

by Ronald Jones, a firearm and toolmark¹ examiner. Jones compared the cases using a microscope and reported on November 4, 2019, that (1) the three cartridge cases recovered at the scene of Wright’s homicide were all fired from the same unknown firearm and (2) that same unknown firearm fired the one cartridge case recovered from the scene of Brison’s shooting.

On November 6, 2019, a federal grand jury returned a nine-count third superseding indictment against Defendant. The first seven counts arose from Defendant’s unlawful possession of a firearm and drug-trafficking activities five years before the shooting. This appeal concerns only the last two counts. Count 8 charged that Defendant violated 18 U.S.C. § 922(g)(1) by being a felon in possession of ammunition—the Blazer 9mm Luger cartridge case recovered from the scene of the January 2019 incident. Count 9, which was added in the third superseding indictment, charged that Defendant violated the same provision in February 2019 by possessing the one Blazer 9mm Luger cartridge case and two Winchester 9mm Luger cartridge cases recovered from the scene of Wright’s homicide.

After the government disclosed that it intended to present firearms-expert testimony to show that the spent cartridge cases recovered from the January and February incidents were fired from the same gun, Defendant filed in March 2020 his

¹ “Toolmarks are generated when a hard object (tool) comes into contact with a relatively softer object. Such toolmarks may occur in the commission of a crime when an instrument such as a screwdriver, crowbar, or wire cutter is used or when the internal parts of a firearm make contact with the brass and lead that comprise ammunition.” Nat’l Rsch. Council, *Strengthening Forensic Science in the United States: A Path Forward* 150 (2009).

“Motion in Limine to Exclude Ballistics Evidence, or, Alternatively, for a *Daubert* Hearing.” R., Vol. I at 106.

Before our discussion of the arguments advanced in Defendant’s motion to exclude the firearms-expert testimony, an overview of firearm toolmark examination is in order. The method of comparing marks left on different pieces of ammunition to determine whether the ammunition was expended from the same firearm has been used for over a century. *See, e.g., Commonwealth v. Best*, 62 N.E. 748, 750 (Mass. 1902) (Holmes, C.J.) (rejecting a challenge to the admission of an expert’s firearms-identification testimony). While advances in science and technology have refined the field of firearm toolmark examination, its object remains the same: “to determine whether ammunition is or is not associated with a *specific* firearm based on toolmarks produced by guns on the ammunition.” President’s Council of Advisors on Sci. and Tech., *Forensic Science in Criminal Courts: Ensuring Scientific Validity of Feature-Comparison Methods* 104 (2016) [hereinafter PCAST Report]. The underlying concept is that different equipment and processes used in manufacturing firearms produce unique marks on the internal parts of the firearm. Firearm examiners theorize that even the same manufacturing tool will produce “microscopically different” marks on consecutively produced firearms, since “[m]anufacturing tools experience wear and abrasion as they cut, scrape, and otherwise shape metal.” Nat’l Rsch. Council, *Strengthening Forensic Science in the United States: A Path Forward* 150 (2009) [hereinafter NRC Report]. Also, the internal parts of the firearm may undergo individualized changes through use. Firearm toolmarks are then produced “when the

internal parts of a firearm make contact with the brass and lead that comprise ammunition.” *Id.* Specifically with respect to cartridge cases (which house the primer, gunpowder, and bullet):²

The brass exterior of cartridge cases receive[s] . . . toolmarks during the process of gun firing: the firing pin dents the soft primer surface at the base of the cartridge to commence firing, the primer area is forced backward by the buildup of gas pressure (so that the texture of the gun’s breech face is impressed on the cartridge), and extractors and ejectors leave marks as they expel used cartridges and cycle in new ammunition.

Id. at 151.

The experts in this case used the Association of Firearms and Tool Mark Examiners (AFTE) method to compare the toolmarks left on the four cartridge cases. *See United States v. Hunt*, 464 F. Supp. 3d 1252, 1256 (W.D. Okla. 2020). Under this method, examiners begin by evaluating class characteristics, “which are features that are permanent and predetermined before manufacture.” PCAST Report at 104. They are “family resemblances which will be present in all weapons of the same make and

² At oral argument Defendant’s appellate counsel suggested for the first time that the district court erred to the extent that it found firearm toolmark examination reliable based on studies involving toolmarks left on bullets (as opposed to cartridge cases). We decline to consider this argument. *See United States v. Malone*, 10 F.4th 1120, 1124–25 (10th Cir. 2021) (“[I]ssues may not be raised for the first time at oral argument.” (internal quotation marks omitted)). We note, however, that while the processes that transfer toolmarks to bullets are not the same as those leaving marks on cartridge cases, *see* NRC Report at 151, the examinations of each “involve[] many of the same concepts,” *United States v. Williams*, 506 F.3d 151, 158 n.4 (2d Cir. 2007); *see also* Alfred Biasotti, John Murdock & Bruce R. Moran, *Introductory Discussion of the Science—The Scientific Methods Applied in Firearms and Toolmark Examination*, in *4 Modern Scientific Evidence: The Law and Science of Expert Testimony* § 34:9 (David L. Faigman et al. eds., 2021) (discussing the general method of toolmark examination).

model.” *United States v. Willock*, 696 F. Supp. 2d 536, 557–58 (D. Md. 2010) (internal quotation marks omitted). For cartridge cases this includes caliber and the shape of the firing pin (which affects the firing pin’s impression on the cartridge). “If the class characteristics are similar, the examination proceeds to identify and compare individual characteristics, such as the striae that arise during firing from a particular gun.” PCAST Report at 104. Examiners can make an “Identification”—that is, conclude that two specimens were derived from the same source (such as a firearm)—when there is:

Agreement 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.

Ass’n of Firearm and Tool Mark Exam’rs, *Glossary* 94 (6th ed. 2013) (defining *Identification*) [hereinafter *AFTE Glossary*]. In arriving at an identification the examiner tries to avoid confusing individual characteristics with subclass characteristics—features produced during manufacture, not by design, that are “consistent among items fabricated by the same tool in the same approximate state of wear.” *Id.* at 121.³ On the opposite end of the spectrum, when there is “[s]ignificant

³ For example, “an imperfection on a rifling tool” can “impart[] similar tool marks on a number of barrels before being modified either through use or refinishing.” Ronald G. Nichols, *Defending the Scientific Foundations of the Firearms and Tool Mark Identification Discipline: Responding to Recent Challenges*, 52 J. Forensic Sci. 586, 587 (2007). An examiner who is not attuned to this possibility or lacks adequate training or experience may make a misidentification of two bullets as having come from the same weapon, when the two weapons that fired

disagreement of discernible class characteristics and/or individual characteristics,” an examiner can determine that the specimens came from different sources—this is known as an “Elimination” conclusion. *Id.* at 94 (defining *Elimination*).

In recent years the field of firearm toolmark examination has been subjected to closer scrutiny as part of a general program to examine the validity of feature-comparison forensic methods used in the courts. We discuss three efforts which were of special importance in the district court’s resolution of Defendant’s pretrial motion: (1) the 2009 NRC Report, (2) a 2014 study by the Ames Laboratory, and (3) the 2016 PCAST Report.

In 2005 Congress authorized the National Academy of Sciences to create a committee to study the needs and practices of the forensic-science community. *See* NRC Report at 1–2. The committee issued a report in 2009 in which it recommended that further research be conducted “to address issues of accuracy, reliability, and validity in the forensic science disciplines.” *Id.* at 22. With respect to firearm toolmark examination, the report found that “[s]ufficient studies have not been done to understand the reliability and repeatability of the methods.” *Id.* at 154. It also criticized the discipline’s “lack of a precisely defined process,” stating that “AFTE has adopted a theory of identification, but it does not provide a specific protocol.”

the different bullets merely share a subclass similarity. *See* Adina Schwartz, *A Systemic Challenge to the Reliability and Admissibility of Firearms and Toolmark Identification*, 6 Colum. Sci. & Tech. L. Rev. 1, 8–10 (2005).

Id. at 155. The NRC Report recognized, however, that firearm toolmark evidence can be valuable:

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

Id. at 154.

In response to the NRC Report’s recommendations, research was conducted to better assess the validity of firearm toolmark examination. We discuss here a study by the Ames Laboratory. *See* 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) [hereinafter Ames Study]. The Ames Study tested 218 firearm examiners by sending each of them 15 sets of four spent cartridge cases. *See* Ames Study at 3–5, 8–10. The researchers generated the spent cartridge cases using 25 Ruger SR9 semiautomatic 9mm handguns.⁴ *See id.* at 5, 9–10. The examiners were to determine whether the fourth cartridge case in each set (the questioned case) came from the same firearm as the three other cartridge cases in that set (the known cases). *See id.* at 4, 10, 15–16. For each examiner the three cartridge cases came from a different firearm than the fourth case in 10 of the 15 sets. *See id.* at 10. The examiners thus conducted 2,180 “true different-source

⁴ While the PCAST Report approved of the Ames Study’s research design, it did note one forensic scientist’s criticism that “the study did not involve *consecutively manufactured* guns.” PCAST Report at 110 n.331.

comparisons” in total: among these, there were 1,421 correct elimination conclusions, 735 inconclusive findings, and 22 erroneous identification conclusions. *Id.* at 16.

The false-positive rate in the Ames Study, excluding inconclusive determinations, was 1.52%. This may overestimate the error rate for firearm toolmark examinations that most law-enforcement agencies and laboratories conduct. Notably, “[a]ll but two of the 22 false identification calls were made by five of the 218 examiners, strongly suggesting that this error probability is not consistent across examiners.” *Id.* And “the study specifically asked participants not to use their laboratory or agency peer review process.” *Id.* at 5. Therefore, the error rate for comparisons that are peer-reviewed could be lower than 1.52%.

In 2015 the President’s Council of Advisors on Science and Technology began investigating “whether there [were] additional steps on the scientific side, beyond those already taken by the Administration in the aftermath of the highly critical 2009 [NRC Report], that could help ensure the validity of forensic evidence used in the Nation’s legal system.” PCAST Report at x. Like its predecessor, the PCAST Report—issued in 2016—said that there is a need for additional studies to verify the principles and methods underlying firearm toolmark examination; it found “that firearms analysis currently falls short of the criteria for foundational validity, because there is only a single appropriately designed study to measure validity and estimate reliability”—the Ames Study.⁵ *Id.* at 112. The PCAST Report also complained that

⁵ According to PCAST, “*Foundational validity* for a forensic-science method requires that it be shown, based on empirical studies, to be *repeatable, reproducible,*

the standard used by toolmark examiners in making an identification is ill-defined, asserting that the theory of identification is “circular”: “The ‘theory’ states that an examiner may conclude that two items have a common origin if their marks are in ‘sufficient agreement,’ where ‘sufficient agreement’ is defined as the examiner being convinced that the items are extremely unlikely to have a different origin.” *Id.* at 104.

But the PCAST Report did not call for the immediate, wholesale exclusion of firearm toolmark evidence from the courts. Instead it took the position that “[w]hether firearms analysis should be deemed admissible based on current evidence is a decision that belongs to the courts.” *Id.* at 112. The PCAST Report’s assessment of this field stands in stark contrast to its assessment of the state of at least one other forensic-science method that involves feature comparison—bitemark analysis. *See id.* at 87 (“[B]itemark analysis does not meet the scientific standards for foundational validity, and is far from meeting such standards.”).

With this background in mind, we now turn to Defendant’s motion to exclude expert testimony regarding firearm toolmark examination. His core argument was that “the Government cannot show that the field of ballistics identification is

and *accurate*, at levels that have been measured and are appropriate to the intended application. . . . It is the *scientific* concept we mean to correspond to the *legal* requirement, in [Federal Rule of Evidence] 702(c), of ‘reliable principles and methods.’” PCAST Report at 4–5. PCAST recognized that “[s]ome methods that have not been shown to be foundationally valid may ultimately be found to be reliable, although significant modifications to the methods may be required to achieve this goal.” *Id.* at 14.

scientifically valid.” R., Vol. I at 109.⁶ After discussing the critiques presented in the PCAST and NRC Reports, he proceeded to argue that exclusion was warranted under Federal Rule of Evidence 702, applying the nonexclusive factors articulated in *Daubert v. Merrell Dow Pharmaceuticals, Inc.*, 509 U.S. 579 (1993), for assessing whether an expert’s methodology is sufficiently reliable for the expert’s opinion to be admissible evidence. The *Daubert* factors are: “(1) whether the [methodology] can be tested; (2) whether it is subject to peer review and publication; (3) the known or potential error rate; (4) the existence and maintenance of standards [to control the methodology’s application]; and (5) the general acceptance in the relevant scientific community.” *United States v. Foust*, 989 F.3d 842, 845 (10th Cir. 2021) (citing *Daubert*, 509 U.S. at 593–94). Defendant argued:

On each *Daubert* factor, ballistics identification evidence falls short. First, the very premise of firearms analysis as a field—the theory of uniqueness of firearms, bullets and cartridge cases—rests on an assumption that has not been properly tested. It comes from a time when bullets were hand-cut and unique, but the field simply has not reckoned with standardization and technological developments in production. . . .

Second, “peer review and publication” is limited. There has only been one black-box study on firearms identification. *See* PCAST Report at 11, 111. This study was not published in a scientific journal and was not subject to peer review or publication. *Id.*

Third, [Defendant] is aware of only one black-box study conducted with respect to error rates. *See [i]d.* at 11. The single study estimated an error rate, but a single study estimating such an error rate is insufficient to determine a known rate of error for the field. . . .

Fourth, there are no uniform standards for controlling the technique’s operation. Rather, an individual makes a subjective

⁶ Defendant’s motion also argued that the government “has not shown that [its] firearm expert, Ron Jones, conducted his examination of the ballistics evidence in a reliable manner.” R., Vol. I at 109. That argument is not renewed on appeal as to Howard Kong, the expert who testified at trial.

determination about an identification, based on his or her own internal compass. There is no required number of points of agreement for making an identification. . . .

Although courts have also previously allowed firearms examiners to testify, the awareness of the limitations of ballistics identification is recent and growing. It does not appear that courts have yet seriously considered all aspects of the field’s development, or tested its reliability since the PCAST report decided it was not foundationally valid from a scientific perspective. . . . Since firearms identification meets none of the *Daubert* criteria, this Court should exclude testimony from the government’s proposed ballistics identification expert.

R., Vol. I at 118–20. Defendant also contended that if the firearms-identification evidence was admitted, (a) the expert should not be permitted to overstate the certainty of his conclusion that the cartridge cases were a match and (b) the jury (1) should be “made aware that the foundational validity of firearms analysis as a field has not been established” and (2) should be informed of the error rates documented in the only “appropriately designed” study. *Id.* at 124.

The government filed a response to Defendant’s motion in which it identified a second firearm expert, Howard Kong, who is a firearm and toolmark examiner with the Bureau of Alcohol, Tobacco, Firearms and Explosives (ATF) Forensic Science Laboratory in California. Kong became involved in the case after images of the two Blazer cartridge cases—one from each incident—were separately uploaded into the ATF’s National Integrated Ballistics Information Network (NIBIN), a database of “three dimensional digital ballistic images of spent shell casings recovered from crime scenes and from crime gun test-fires” that “can automatically generate a list of potential matches,” purportedly with a “very high level of accuracy.” Erin Aslan,

Crime Gun Intelligence Centers: Using Technology and Intelligence as a Lead Generator to Identify Trigger-Pullers and Focus Enforcement and Prevention Efforts, DOJ J. Fed. L. & Prac., Nov. 2018, at 49, 53. The NIBIN system identified a potential link between the two Blazer cartridge cases; the images were then reviewed by a correlation review specialist, a peer reviewer, and an ATF firearm examiner before a NIBIN lead was generated, advising that further investigation was warranted. All four spent cartridge cases were then sent to the ATF laboratory, where Kong—who had been advised only that there was a NIBIN lead—conducted a microscopic comparison and concluded that “the probability that these cartridge cases were fired in a different firearm is so small that it is negligible.” R., Vol. I at 216. A peer reviewer reached the same conclusion.

The district court denied Defendant’s motion to exclude, without a hearing, in a 17-page order that applied the *Daubert* factors to analyze the validity of firearm toolmark examination as a field and considered whether Jones and Kong reliably applied firearm-examination methods in this case. The court observed that the use of firearm toolmark examination in court is far from novel and that “no federal court has deemed such evidence wholly inadmissible.” *Hunt*, 464 F. Supp. 3d at 1256. It then analyzed the five *Daubert* factors. *See id.* at 1256–60.

First, the district court found that “the theory of firearm toolmark identification can be and has been tested.” *Hunt*, 464 F. Supp. 3d at 1257. It cited a collection of studies compiled by the AFTE and noted that Defendant had presented no authority to the contrary. *See id.* at 1256–57.

Second, the court found that *Daubert*'s peer-review factor also favored admission, citing the AFTE Journal (a publication for firearm toolmark research), the PCAST and NRC Reports, and various studies that have been conducted on firearm toolmark examination. *See id.* at 1257–58.

Third, the court found that the low error rate favored admissibility, citing the 1.52% false-positive rate reported in the Ames Study⁷ and “a Miami-Dade Study that reported a potential error rate of less than 1.2%.” *Id.* at 1258. The court observed that “[o]ther federal courts examining the AFTE method’s rate of error have likewise found it to be low” and noted that Defendant had not “introduce[d] any contradictory studies.” *Id.*

On the other hand, the district court found that the “standards that control the [methodology’s application]” factor weighed against admissibility because the

⁷ In Defendant’s view the real error rate of the Ames Study was roughly 35%. His argument is that inconclusive determinations should be considered errors because “the samples were controlled to ensure that they bore sufficient markings and that these markings were [sic] not disturbed by environmental factors.” *Aplt. Br.* at 27. He cites a portion of the Ames Study that mentions that the fired cartridge cases were collected in a brass catcher and that cases that fell out of the catcher were discarded. *See Ames Study* at 11–12. But this is weak support for his proposition, considering that the Ames Study did not “prescreen[] the quality of samples provided to the participants” and actually sought to collect data on how many samples “had marks that were suitable for comparison.” *Id.* at 4. In any event, we find persuasive the government’s argument that the Ames Study’s false-positive rate furnishes the relevant error rate. That is because “a false positive identification . . . is the type of error that could lead to a conviction premised on faulty evidence.” *United States v. Harris*, 502 F. Supp. 3d 28, 39 (D.D.C. 2020). There is no harm to the defendant if a toolmark examiner makes an inconclusive finding, and Defendant presents no evidence to support his speculation that examiners will feel pressured to render conclusive opinions in the trial setting.

general AFTE method is “subjective in nature.” *Id.* at 1259 (internal quotation marks omitted). The district court concluded, however, that “the subjectivity of a methodology is not fatal under Rule 702 and *Daubert.*” *Id.* at 1260 (internal quotation marks omitted).

Finally, with respect to the general-acceptance factor, the district court found that “the AFTE method used by the Government’s expert here[] is the field’s established standard” and that the lack of universal acceptance—citing the NRC and PCAST Reports—did not preclude admission in court. *Id.* at 1259–60 (internal quotation marks omitted).

Despite the court’s weighing of the factors in favor of admissibility, it restricted how the expert’s opinion could be presented. It required that the testimony adhere to certain limitations set forth in guidance issued by the Department of Justice.⁸ *See id.* at 1261–62. The experts would need to “refrain from expressing their

⁸ The following limitations are set forth in the Department of Justice’s “Uniform Language for Testimony and Reports for the Forensic Firearms/Toolmarks Discipline – Pattern Match Examination,” R., Vol. I at 295:

- An examiner shall not assert that two toolmarks originated from the same source to the exclusion of all other sources. This may wrongly imply that a ‘source identification’ conclusion is based upon a statistically-derived or verified measurement or an actual comparison to all other toolmarks in the world, rather than an examiner’s expert opinion.
- 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.

findings in terms of absolute certainty, and they [would] not state or imply that a particular bullet or shell casing could only have been discharged from a particular firearm to the exclusion of all other firearms in the world.” *Id.* at 1261. The court permitted the experts to testify only “that their conclusions were reached to a reasonable degree of ballistic certainty, a reasonable degree of certainty in the field of firearm toolmark identification, or any other version of that standard.”⁹ *Id.* at 1262.

-
- 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 proffered conclusion. 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 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.

Id. at 297. Defendant concedes that “Kong was able to follow the letter of the [district court’s] order” limiting the testimony but complains that his testimony nonetheless “provided the false sense of certainty that the order was attempting to avoid.” *Aplt. Br.* at 33. Defense counsel did not, however, raise any objections during Kong’s testimony at trial.

⁹ Federal Rule of Evidence 702(d) currently requires that the proponent of expert testimony show that “the expert has reliably applied the principles and methods to the facts of the case.” The Advisory Committee on Evidence Rules has proposed amending Rule 702(d) to require a showing that “the *expert’s opinion reflects a reliable application of the principles and methods to the facts of the case.*” Committee on Rules of Practice and Procedure, Judicial Conference of the United States, *Proposed Amendments to the Federal Rules of Appellate, Bankruptcy, Civil, and Criminal Procedure, and the Federal Rules of Evidence* 308–09 (2021) (emphasis added). The committee note explains that “[f]orensic experts should avoid assertions of absolute or one hundred percent certainty—or to a reasonable degree of scientific certainty—if the methodology is subjective and thus potentially subject to error.” *Id.* at 311.

As relevant here, Defendant proceeded to trial on Counts 8 and 9. Kong testified at Defendant's trial, but Jones did not. Kong outlined his training and experience, which included earning a bachelor's degree in materials engineering, completing a one-year course at the ATF Firearms Examiner Academy, and receiving six to eight months of in-house training. Before he became an official firearms examiner, he was required to pass a competency test. In 2009, after having acquired five years of experience, he passed an examination required for certification by AFTE. AFTE certification lasts for five years but is subject to passing yearly proficiency tests. He has maintained his certification and was recertified in 2014 and 2019. Kong testified that he had 18 years of experience in firearms examination and that he had handled an average of 50 cases a year. Also, he has toured the facilities of over a dozen firearms manufacturers.

Consistent with his written report (which had been submitted to Defendant and the court in response to the pretrial *Daubert* motion), Kong opined that the four spent cartridge cases found at the January and February 2019 crime scenes were fired from the same gun. The class characteristics of the cartridge cases "were all similar." R., Vol. III at 467. With respect to individual characteristics, he explained that he had examined the striated firing-pin aperture shear marks that were created when the "head" of each of the four cartridge cases—that is, "[t]he base of the cartridge case which contains the primer," *AFTE Glossary* at 32—scraped against the breech face of

the firearm upon discharge.¹⁰ Using a comparison microscope to look at the cartridges side by side, he found that the striations in the shear marks were in “sufficient agreement . . . to identify them as having been fired from the same gun.” R., Vol. III at 467. He also found that the firing-pin impressions—marks left by the firing pin upon striking the cartridge cases—were in “excellent agreement.” *Id.* at 468.

In discussing the striated firing-pin aperture shear marks, Kong said that there were “somewhere in the neighborhood of 15 to 20 . . . consecutively matching stria[e].” *Id.* Earlier in his testimony he had mentioned a study conducted by Alfred Biasotti and John Murdock that examined the striated marks left on bullets fired from different firearms and concluded that the best agreement between two different-source bullets was “maybe three or four consecutive matching striations,” that is, lines that are “consecutive and right next to each other.” *Id.* at 462. According to Kong, the study’s authors concluded that “if you have six consecutive stria[e], then that would signify a match.” *Id.* He thus opined that this was “not a borderline case in terms of identification.” *Id.* at 468. He said that he was able to reach his conclusion to “a reasonable degree of certainty within the firearms examination field.” *Id.*

¹⁰ Kong described the breech face as “the backstop or the area that contacts the base of the cartridge case. . . . [W]hen the cartridge fires, it pushes the case back against this breech face and picks up marks.” R., Vol. III at 453. The firing-pin aperture is “[t]he hole in the breech face of a firearm through which the firing pin protrudes.” *AFTE Glossary* at 52. Aperture shear marks are “[s]triated marks caused by the rough edges of the firing pin aperture scraping the primer metal during unlocking of the breech.” *Id.* at 52–53.

Kong also testified that, based on differences in class characteristics, the recovered cartridge casings could not have been fired by two firearms he was asked to examine, firearms that belonged to two of the men who accompanied Defendant during the January 2019 incident. The testimony at trial indicated that Defendant was the only other person—of the four men who confronted Brison—who could have carried a gun during that incident. Defense counsel did not conduct any cross-examination of Kong. Defendant was convicted on all counts that went to trial and sentenced to 960 months in prison.

II. DISCUSSION

We review for abuse of discretion the district court’s application of *Daubert* in denying Defendant’s motion to exclude the firearms-examination evidence. *See Foust*, 989 F.3d at 845. “We give the district court substantial deference, reversing only when its ruling was arbitrary, capricious, whimsical, or manifestly unreasonable or when it made a clear error of judgment or exceeded the bounds of permissible choice in the circumstances.” *Id.* (internal quotation marks omitted).

But “we review de novo the question of whether the district court applied the proper standard [in admitting an expert’s testimony] and actually performed its gatekeeper role in the first instance.” *Dodge v. Cotter Corp.*, 328 F.3d 1212, 1223 (10th Cir. 2003). This is because “[w]hile the district court has discretion in the *manner* in which it conducts its *Daubert* analysis, there is no discretion regarding the actual *performance* of the gatekeeper function.” *Goebel v. Denver & Rio Grande W. R.R. Co.*, 215 F.3d 1083, 1087 (10th Cir. 2000) (emphasis in original). Thus, our de

novo review is limited to whether the district court properly followed the *Daubert* framework and performed an adequate inquiry into the relevance and reliability of the expert testimony. See *Kumho Tire Co. v. Carmichael*, 526 U.S. 137, 147 (1999); *Dodge*, 328 F.3d at 1221. “For purposes of appellate review, a natural requirement of the gatekeeping function is the creation of a sufficiently developed record in order to allow a determination of whether the district court properly applied the relevant law.” *Adamscheck v. Am. Fam. Mut. Ins. Co.*, 818 F.3d 576, 586 (10th Cir. 2016) (brackets and internal quotation marks omitted). And the trial court is required to “reply in some meaningful way to the *Daubert* concerns the objector has raised.” *StorageCraft Tech. Co. v. Kirby*, 744 F.3d 1183, 1190 (10th Cir. 2014).

A. The District Court’s Gatekeeping Role

Defendant does not dispute that the district court cited “the [*Daubert*] factors and looked to them exclusively.” Aplt. Br. at 16. But he contends that the district court “procedurally erred in its application of *Daubert*” because it failed to conduct a diligent assessment of the available empirical evidence before making its ruling. *Id.* at 14. He faults the district court for taking “a short-cut around *Daubert* by relying on prior judicial decisions instead of conducting a meaningful review of the science.”

*Id.*¹¹

¹¹ Defendant appears to argue that this critical assessment should have included a “probing inquiry,” Aplt. Br. at 26, into the methodologies used in the Miami-Dade and Ames Studies, and a thorough analysis of the peer-review procedure and publication practices of the ATFE Journal. But Defendant did not raise these specific concerns in his pretrial brief. As we have explained, the district court’s gatekeeping function is flexible, “requiring the court to focus its attention on the specific

Defendant finds some support in the beginning of the district court’s analysis:

The use of this type of firearm toolmark identification in criminal trials is “hardly novel.” *United States v. Taylor*, 663 F. Supp. 2d 1170, 1175 (D.N.M. 2009). “For decades ... admission of the type of firearm identification testimony challenged by the defendant[] has been semi-automatic” *United States v. Monteiro*, 407 F. Supp. 2d 351, 364 (D. Mass. 2006); *see also, e.g., United States v. Hicks*, 389 F.3d 514 (5th Cir. 2004); *United States v. Johnson*, 875 F.3d 1265, 1281 (9th Cir. 2017). Indeed, no federal court has deemed such evidence wholly inadmissible. *See United States v. Romero-Lobato*, 379 F. Supp. 3d 1111, 1117 (D. Nev. 2019).

Hunt, 464 F. Supp. 3d at 1256. If this was the court’s only analysis, and the court had rested its decision solely on other courts admitting firearm toolmark identification evidence, it may well have failed to perform its gatekeeping duty. But the district court continued, stating that “because of the seriousness of the criticisms launched against the methodology underlying firearms identification by Defendant in this case, the Court will carefully assess the reliability of this methodology, using *Daubert* as a guide.” *Id.* at 1256. The district court then worked through the *Daubert* factors, while considering the arguments presented in Defendant’s pretrial brief. *See id.* at 1256–60. There is nothing improper in a court adopting the reasoning of a prior court. And the decision primarily relied upon by the district court to support its findings, *United States v. Romero-Lobato*, discussed the PCAST and NRC reports in depth. *See* 379 F. Supp. 3d 1111, 1117–18 (D. Nev. 2019), *aff’d in relevant part*, No. 20-10280, 2022 WL 2387214, at *1 (9th Cir. July 1, 2022) (unpublished). Although Defendant argues

factors implicated by the circumstances at hand.” *StorageCraft*, 744 F.3d at 1190. The district court is not required to discuss issues that are not raised by the parties.

that “[t]he district court effectively took an end-run around the PCAST Report by relying on other decisions that either could not or did not consider the report,” Aplt. Reply Br. at 4, he ignores the district court’s discussion of the PCAST Report’s critiques that were raised by Defendant in his pretrial briefing, *see Hunt*, 464 F. Supp. 3d at 1257–58. Defendant may disagree with the manner in which the district court fulfilled its gatekeeping responsibilities, but because the district court analyzed the *Daubert* factors and addressed Defendant’s arguments, we conclude that the district court’s gatekeeping role was performed.

B. Admissibility of the Expert’s Testimony

Defendant argues that the district court abused its discretion in allowing Kong to testify regarding his firearm toolmark examination. He claims that “firearm toolmark examination methods are subjective, unproven, and not subject to meaningful review or acceptance outside the insular community of firearm toolmark examiners.” Aplt. Br. at 15. Although he challenges the reliability of the AFTE method, he does not challenge Kong’s credentials or whether he reliably applied the methodology to the facts of this case.

In reviewing the district court’s ruling, we are not limited to its exposition supporting the ruling. Even if the exposition may be deficient in some respects, any shortcoming may be harmless error if the record contains the necessary support. As stated in *StorageCraft*, “If . . . it is readily apparent from the record that the expert testimony *was* admissible, it would be pointless to require a new trial at which the very same evidence can and will be presented again.” 744 F.3d at 1191. In particular,

we may evaluate the ruling in light of evidence presented at trial that was not presented at the *Daubert* hearing. *See Kinser v. Gehl Co.*, 184 F.3d 1259, 1271 (10th Cir. 1998) (In reviewing the admissibility of expert testimony, “we do not think it necessary to confine our review to the materials accompanying the *Daubert* hearing request. Rather, we believe we may look at the entire record, including testimony presented at trial.”), *abrogated in part on other grounds by Weisgram v. Marley Co.*, 528 U.S. 440 (2000); *see also StorageCraft*, 744 F.3d at 1191–92 (in response to an argument that the expert made an assumption unsupported by the evidence, the court held that any error was harmless because the assumption was supported by party’s deposition); *cf. United States v. Harrison*, 296 F.3d 994, 1003 (10th Cir. 2002) (“[T]o affirm the district court’s ruling, we may decide to consider all the evidence at trial, including evidence not presented at the hearing on the motion *in limine* [to admit evidence].”); *United States v. Corral*, 970 F.2d 719, 723 (10th Cir. 1992) (“In evaluating the correctness of the district court’s rulings, the appellate court may consider the entire record developed from the trial even though such evidence may not have been presented during the suppression hearing.”).

Our concern is the breadth of the district court’s ruling. It can be read as upholding AFTE methodology in general. This would hardly be a remarkable ruling. Both before and after the 2016 publication of the PCAST Report, other circuits have affirmed decisions to admit expert testimony grounded in the AFTE method. *See United States v. Brown*, 973 F.3d 667, 702–04 (7th Cir. 2020); *United States v. Johnson*, 875 F.3d 1265, 1281 (9th Cir. 2017); *United States v. Williams*, 506 F.3d

151, 157–62 (2d Cir. 2007). But in light of the critiques expressed in the PCAST and NRC Reports, we think courts should be cautious, and our holding should go no further than necessary. “Our task is not to determine the admissibility or inadmissibility of [firearm toolmark examination] for all cases but merely to decide whether, on this record, the district judge in this case made a permissible choice in exercising [his] discretion to admit the expert testimony.” *United States v. Baines*, 573 F.3d 979, 989 (10th Cir. 2009).

We therefore restrict our consideration to the specific methodology described by expert Kong at trial. To determine that all four cartridges were fired by the same weapon, Kong employed the consecutive-matching-striae (CMS) method to three-dimensional images of the cartridges. As we explain below, the CMS method has impressive empirical support that would plainly permit a court to find it reliable. The rub in this case is that the district court’s Order expressing its ruling on the *Daubert* motion does not describe that method or the technical literature supporting it. But that is hardly surprising. Until Kong testified, nothing in the record would have informed the court about the specific identification method used by Kong.¹² Although the technical literature referenced by the government in its *Daubert* brief and by the court in its Order says a good deal about the CMS method, there was no reason for

¹² The government first identified Kong as an expert in its response to Defendant’s motion to exclude. Defendant filed no reply challenging Kong’s qualifications or expertise. At trial Hunt did not cross-examine Kong or raise any objection to his testimony. He thereby forfeited any challenge that might derive from the failure to disclose this methodology before trial. Nor has Hunt raised such a challenge on appeal.

the Order to focus on that method as opposed to general AFTE methods. But if the court had been called on at trial to opine on this specific subset of AFTE methodology, there can be little doubt that the court, having already declared that the general AFTE methodology is reliable, would have found that the CMS method, with its extensive empirical support, is a fortiori reliable. Hence, any error in the court's not specifically addressing the CMS method is harmless. We now turn to a discussion of that method and how it would affect the district court's *Daubert* analysis.

In applying the CMS method, the examiner aligns the objects being compared and counts the number of consecutive striae where the width, morphology, and relative position match exactly. Kong testified that finding six or more consecutive striae that match is sufficient to determine that two cartridges were fired from the same weapon. He also said that each additional consecutive matching stria further reduces the likelihood that two cartridges were fired from different weapons. Thus, because the four cartridge cases here had 15 to 20 consecutive matching striae, it was “not a borderline case in terms of identification.” R., Vol. III at 468.

Kong said that the criterion he used was based on

a study done some time ago by Biasotti and Murdock. And what they did was they looked at the best agreement and bullets that were fired from different firearms, and the best they came up with was, like, maybe three or four consecutive matching striations, and that's the absolute best. So what they did was they proposed that if you have six consecutive strias, then that would signify a match.

Id. at 462. He further testified that if two bullets (or cartridges) fired from different guns had striae that satisfied the criterion for identifying the bullets as coming from

the same weapon, firearms examiners would learn of the discovery, yet he had never heard of such a discovery.

The criterion used by Kong is supported by the research of Biasotti and Murdock that he mentioned in his testimony and by several studies referenced in the government's pretrial *Daubert* brief and in the district court's Order. Biasotti's original study, first published in 1959, examined 244 bullets and analyzed 1,200 known match comparisons and 1,080 known nonmatch comparisons. It found empirical support for the proposition that a significant number of consecutive matching striae will appear only on bullets fired from the same gun.¹³ See Biasotti, *A Statistical Study of the Individual Characteristics*, *supra*, at 35–36; Alfred Biasotti, John Murdock & Bruce R. Moran, *Development of Objective Criteria for Identification*, in *4 Modern Scientific Evidence: The Law and Science of Expert Testimony* § 34:13, at 1013–18 (David L. Faigman et al. eds., 2021). Biasotti and Murdock continued to research consecutive matching striae in toolmarks, and in 1997 they formulated their “conservative quantitative criteria for identification” (the Biasotti-Murdock criteria). Jerry Miller, *An Examination of the Application of the Conservative Criteria for Identification of Striated Toolmarks Using Bullets Fired from Ten Consecutively Rifled Barrels*, 33 AFTE J. 125, 126 (2001); see Biasotti,

¹³ Before this research some experts looked to the total percent of matching striae (regardless of whether they were consecutive) between the two bullets to determine a match. But Biasotti concluded from his data that the total percent of matching striae was an unreliable basis for identification. See Alfred Biasotti, *A Statistical Study of the Individual Characteristics of Fired Bullets*, 4 J. Forensic Sci. 34, 37–39 (1959).

Murdock & Moran, *Development of Objective Criteria, supra*, at 1018–20. Under their criteria a match is found for three-dimensional striae “when at least two different groups of at least three consecutive matching striae appear in the same relative position, or one group of six consecutive matching striae are in agreement.”¹⁴

Michael Neel & Major Wells, *A Comprehensive Statistical Analysis of Striated Tool Mark Examinations Part 1: Comparing Known Matches and Known Non-Matches*, 39 AFTE J. 176, 177 (2007). But to apply the criteria, the toolmark examiner must rule out the influence of subclass characteristics.¹⁵

¹⁴ Biasotti and Murdock also formulated a criterion for examining two-dimensional toolmarks. Two-dimensional toolmarks, also sometimes called *lines*, include “[a]ny impressed or striated toolmark that lacks apparent depth.” Neel & Wells, *A Comprehensive Statistical Analysis of Striated Tool Mark Examinations, supra*, at 177.) Under the two-dimensional criterion, a match occurs “when at least two groups of at least five consecutive matching striae appear in the same relative position, or one group of eight consecutive matching striae are in agreement in an evidence toolmark compared to a test toolmark.” *Id.* Biasotti’s original research examined two-dimensional marks.

¹⁵ Kong’s notes from his examination of the cartridge cases, which were submitted to the court, discussed his elimination of subclass characteristics:

Aperture shear marks: The aperture shapes are similar to a “teardrop[,]” which means there is a ramp at the 6 o’clock position. The aperture shear marks had excellent correspondence, and the agreement was sufficient for identification. Aperture shear marks were generated when the primer of the cartridge case rubbed on the edge of the breech face at the ramp. The edge is the intersecting of these two differently machined surfaces and will not have subclass potential. The ramp was produced by a rotating cutting tool to a tilted breech, therefore producing toolmarks that are perpendicular to the [cartridge cases’] movement during firing; so no potential subclass marks are from the ramp.

Firing pin marks: Excellent correspondence of a series of marks observed on the firing pin (FP) impression. A line was visible running from the 6 to 12 o’clock in the center of the FP impression, and

The Biasotti-Murdock criteria, the basis of the CMS method, are supported in several ways by technical papers referenced in the government’s *Daubert* brief and the district court’s Order. One 2007 empirical study looked at “4188 striated toolmark comparisons from a variety of sources”¹⁶ and found no known nonmatches that met the Biasotti-Murdock criteria. *Id.* at 179, 180. In particular, the study found no more than four consecutive matching three-dimensional striae in a known nonmatch—two less than the six required under the Biasotti-Murdock criterion. The study pointed out that “[t]here has been over 50 years of research regarding CMS run counts, beginning with Biasotti’s thesis in 1955, and no documented 2D or 3D comparisons have been shown to contradict the original criteria set forth by Biasotti and Murdock.” *Id.* at 190.

The Biasotti-Murdock criteria are also supported by a theoretical analysis that calculated the probabilities of consecutive matches of two-dimensional striae

appeared to be a mold line from castings or metal injection molding (MIM) parts. Mold surfaces can have features carryover from part to part. However, the firing pin appeared to have a defect that produced the observed corresponding marks, and such a defect could be unique, but can have subclass potential. If the mold has a defect, then there is subclass.

R., Vol. I at 272.

¹⁶ The sources for this study included “fired cartridge cases, fired bullets, sandpaper of various grit sizes (60, 22, 320) used to scratch emulsion based film, chisels slid across lead foil, photomicrographs of plastic replicas taken from consecutively rifled barrels, fired bullets from cut sections of a Thompson Contender barrel, and fired bullets from consecutively rifled Bar-Sto barrels.” Neel & Wells, *A Comprehensive Statistical Analysis of Striated Tool Mark Examinations, supra*, at 179.

occurring by chance. See David Howitt, et al., *A Calculation of the Theoretical Significance of Matched Bullets*, 53 J. Forensic Sci. 868 (2008). The researchers found a “close similarity” between Biasotti’s 1959 experimental observations and their calculated probabilities. *Id.* 872–73. Based on their calculations, the probability of the random creation of eight consecutive matching two-dimensional striae (the Biasotti-Murdock two-dimensional criterion) would be less than one in a hundred million.

In addition, the Biasotti-Murdock criteria have been successfully tested in an automated identification system. See Wei Chu, et al., *Automatic Identification of Bullet Signatures Based on Consecutive Catching Striae (CMS) Criteria*, 231 Forensic Sci. Int’l 137 (2013). The system minimized subjective factors by using a computer program objectively applying the three-dimensional Biasotti-Murdock criterion. It correctly identified 29 of 30 matching bullet pairs from the “unknown” test set and found no false positives in 12,960 known nonmatch comparisons.

To be sure, one can find literature criticizing the CMS method. Mr. Hunt’s reply brief cites a 2005 article asserting that striae counting is inherently subjective and that “the CMS approach fails to place firearms and toolmark identification on adequate statistical empirical foundations.” Schwartz, *A Systemic Challenge*, *supra*, at 21. But even that article recognizes advantages of CMS identification as compared to the general AFTE methodology. *Id.* at 15 (“CMS differs from and is scientifically superior to the subjective approach because it is interpretable in a way that is compatible with the probabilistic nature of identity claims.”). And more importantly,

articles discussed above that were published after 2005 undercut Schwartz's concerns. The article by Chu, et al., shows that CMS can be applied objectively; and the article by Howitt, et al., provides theoretical statistical support for CMS. We also think it significant (1) that the CMS method is not mentioned in the PCAST report, which questions only AFTE methods in general and (2) that the 2009 NRC Report suggests that its criticism of the lack of studies supporting firearms and toolmark identification in general may not apply to CMS: "Recent research has attempted to develop a statistical foundation for assessing the likelihood that more than one tool could have made specific marks by assessing consecutive matching striae, but this approach is used in a minority of cases." NRC Report at 154 n.63.

We now turn to an examination of the district court's *Daubert* analysis as supplemented by specific information on the reliability of the CMS method. Regarding the first *Daubert* factor, the district court found that "the theory of firearm toolmark identification can be and has been tested." *Hunt*, 464 F. Supp. 3d at 1257. Discussion of the CMS method would only have reinforced that conclusion. By providing quantitative criteria for identification, the CMS method is readily testable and has been tested. As discussed above, the CMS method has been empirically tested for over 60 years and no nonmatches have been found that meet the Biasotti-Murdock criteria. See Neel & Wells, *A Comprehensive Statistical Analysis of Striated Tool Mark Examinations*, *supra*, at 190; see also Biasotti, Murdock & Moran, *Development of Objective Criteria*, *supra*, at 1021–23 & n.26 (summarizing the empirical studies evaluating Biasotti-Murdock criteria and concluding that "[n]o

known non-matching . . . toolmarks were found . . . which exhibited agreement in excess of the proposed Biasotti-Murdock criteria.”).

The district court also determined that the second *Daubert* factor—whether the technique has been subjected to peer review and publication—weighed in favor of admissibility. *See Hunt*, 464 F. Supp. 3d at 1257–58. Again, that determination finds particular support with respect to the CMS method. Biasotti’s original 1959 article was published in the *Journal of Forensic Science*, which is peer-reviewed. *See Biasotti, Murdock & Moran, Development of Objective Criteria, supra*, at 1013 & n.5. And the other studies discussed above were each published in the *AFTE Journal*, the *Journal of Forensic Science*, or *Forensic Science International*. The studies supporting the CMS methodology were subject to the same peer-review process as the articles addressed by the district court in its Order. This factor would only be strengthened in support of reliability.

The third *Daubert* factor is whether the technique has a known or potential rate of error. The district court, noting the error rates from the Ames Study (which, as here, involved cartridge-case comparisons) and the Miami-Dade Study, weighed this factor in favor of admissibility. *See Hunt*, 464 F. Supp. 3d at 1258.¹⁷ For the Biasotti-

¹⁷ We add that at least one “post-PCAST Report study . . . followed the PCAST recommended black-box model and found that of 1512 possible identifications tested, firearms examiners correctly identified 1508 casings to the firearm from which the casing was fired.” *Harris*, 502 F. Supp. 3d at 38 (citing Mark A. Keisler et al., *Isolated Pairs Research Study*, 50 *AFTE J.* 56, 58 (2018)). Notably, “[n]o false positive . . . results were reported.” Keisler et al., *Isolated Pairs Research Study, supra*, at 57.

Murdock criteria, there is not a known error rate from empirical testing because no false match has been found. But the theoretical analysis discussed above suggests that the probability of two random bullets satisfying the Biasotti-Murdock criteria would be infinitesimal. Again, consideration of the CMS method would only have increased the support of this factor in favor of reliability.

The district court weighed the fourth factor—whether there are standards that control the technique’s operation—against admissibility. *See Hunt*, 464 F. Supp. 3d at 1259. But even if the court’s assessment is correct with respect to toolmark examination in general, the Biasotti-Murdock criteria provide specific objective standards that control the application of the CMS method. *See Romero-Lobato*, 379 F. Supp. 3d at 1121 (“The CMS method, standing alone, qualifies as an objective standard under *Daubert*.”). Although the determination of subclass features and deciding whether two striae match may involve some subjectivity, the toolmark examiner’s discretion is constrained: “CMS is defined as striated markings that ‘line up’ exactly (close doesn’t count) with one another without a break or dissimilarity in between them.” Biasotti, Murdock & Moran, *Development of Objective Criteria*, *supra*, at 1015–16 n.8.

Finally, the district court found that the general-acceptance factor favored admission because the AFTE method is widely used by firearms examiners. *See Hunt*, 464 F. Supp. 3d at 1259–60; *see also Brown*, 973 F.3d at 704 (district court observed that “firearm and toolmark analysis is widely accepted beyond the judicial system”). Like the other factors, consideration of the CMS method would have

further supported this conclusion. Although the CMS method might not be “routinely used by firearm examiners across the country,” *Romero-Lobato*, 379 F. Supp. 3d at 1122, this may be because “many firearms examiners consider [the Biasotti-Murdock criteria] to be too stringent,” *United States v. Taylor*, 663 F. Supp. 2d 1170, 1178 (D.N.M. 2009). This feature of the CMS method—missing some identifications through the application of stringent criteria that minimize the possibility of a false positive—enhances the reliability of CMS-derived identifications. There is no reason to doubt that CMS identifications are generally accepted, at least within the community of firearms examiners. *See Foust*, 989 F.3d at 846 (“Although . . . acceptance by unbiased experts is always better, that does not mean this factor cannot support admission.”); *Baines*, 573 F.3d at 991 (similar).¹⁸

We conclude that even if the district court’s opinion is inadequate to support the AFTE methodology for firearm toolmark examination in general, any shortcoming was harmless because the specific CMS method relied on by Kong has compelling support in his testimony and the technical literature referenced by the district court. And Kong’s opinion in this case seems particularly resistant to criticism given the large number of consecutive matching striae and his expertise and experience in applying the methodology.

¹⁸ We note that one of the most critical judicial assessments of firearm toolmark identification—repeatedly cited by Defendant—did not consider the CMS method. *See generally United States v. Tibbs*, No. 2016 CF1 19431, 2019 WL 4359486 (D.C. Super. Ct. Sept. 5, 2019).

C. Response to the Concurrence

The Concurrence complains that the majority opinion improperly goes outside the record and violates the principle of party presentation. Both complaints are unwarranted.

First, the record. The record consists of those documents (broadly construing the term to include electronically stored matters, such as video or audio recordings) filed with the court. The scientific and technical papers referenced in the majority opinion that the concurrence complains were not in the record were summarized on a website referenced in the district court's opinion and hyperlinked in the government's memorandum filed with the court for the *Daubert* hearing. The only relevance of the referenced website was the studies it collected. Under any reasonable notion of *record*, those papers were part of the district-court record.¹⁹ They would surely have been part of the record if they were attached as hard copies to the government's brief. Is such useless printing to be required by parties in the future? What purpose is advanced by not considering those papers as part of the record?

The Concurrence suggests that maybe the papers were part of the record but only for a limited purpose. It says that the reference to the studies by the government was just to establish that such studies existed, "not their quality or content."

¹⁹ That was certainly the view of defense counsel at oral argument, who described the website articles as "evidence." A few seconds into his argument, counsel noted the citation to the website in the district court's opinion and suggested that the opinion did not contain a sufficient analysis of the studies themselves, arguing that "[t]he court wasn't conducting a review of the evidence by merely looking at that website." Oral Argument at 1:22–1:27.

Concurrence at 1. But surely a court's role in performing a *Daubert* analysis is not limited to looking at titles and counting how many there are. What sort of a gatekeeper is that? The court must examine contested studies to protect the system from junk science and self-serving ipse dixits by advocates for particular theories. The opposing party has every right, and every incentive, to challenge the quality of the studies proffered by the offering party and the relevance of their content to the issues before the court. Here, Defendant had retained his own ballistics expert, although the expert was never called to testify. The fact that the opposing party made no such challenge hardly means that the quality and content of the studies are not part of the record. A great deal of the record in every case goes unchallenged. If the district court had in fact read some of the referenced papers to assist it in ruling on the *Daubert* motion (which is not entirely implausible since the district-court opinion notes that the court visited the site two weeks before filing the opinion, *see* 464 F. Supp.3d at 1257), would we be required to reverse the court's decision because it went "outside the record" in reaching its decision?

Insisting on a cramped notion of what is part of the record is particularly inappropriate in the present context. The studies at issue do not address adjudicative facts peculiar to the specific case before us. They concern legislative facts that are applicable to a great many, and wide range of, cases; the reliability of firearm toolmark examination must be assessed to determine the admissibility of evidence of such examinations. When the resolution of a dispute turns on legislative facts, courts regularly relax the restrictions on judicial inquiry. For example, while appellate

courts can take judicial notice of adjudicative facts only in limited circumstances, *see* Fed. R. Evid. 201; *Winzler v. Toyota Motor Sales U.S.A., Inc.*, 681 F.3d 1208, 1212–13 (10th Cir. 2012) (Gorsuch, J.) (appellate court may take judicial notice of facts “at any stage of the proceeding if the facts are not subject to reasonable dispute” (internal quotation marks omitted)), there are no such strict limits with respect to legislative facts, *see* Fed. R. Evid. 201 advisory committee note (“In determining the content or applicability of a rule of domestic law, the judge is unrestricted in his investigation and conclusion. He may reject the propositions of either party or both parties. He may consult the sources of pertinent data to which they refer, or he may refuse to do so. He may make an independent search for persuasive data or rest content with what he has or what the parties present.” (quoting Edmund M. Morgan, *Judicial Notice*, 57 Harv. L. Rev. 269, 260–71 (1944))); Edward K. Cheng, *Independent Judicial Research in the Daubert Age*, 56 Duke L.J. 1263, 1293 (2007) [hereinafter Cheng] (“If one takes the Advisory Committee’s adoption of the Morgan view seriously, this conclusion [that the scientific facts used for *Daubert* determinations should be treated as legislative facts] means that the Federal Rules free judges to do independent research in the Daubert context.”). As noted by Prof. Frederick Schauer in *The Decline of “The Record”: A Comment on Posner*, 51 Duq. L. Rev. 51 (2013) [hereinafter Schauer], courts sometimes rely on nothing more than the personal intuitions of its members (as when it selected the burden of persuasion for libel actions based on how it thought the press would react to different standards, *see New York Times v. Sullivan*, 376 U.S. 254, 282 (1964); Schauer at 57–58, 63),

and they sometimes rely on their review of sources never presented to the court (as with the per curiam majority in *Bush v. Gore*, 531 U.S. 98, 103 (2000), and Justices Stevens and Breyer on multiple occasions, including for the court in *Kumho Tire Co. v. Carmichael*, 526 U.S. 137, 142–43 (1999); *see* Schauer at 51 n.3, 56). The propriety of this independent research is somewhat controversial, *see, e.g.*, Schauer at 51–52; Cheng at 1280–84, but when the highest court in the land uses non-record materials to resolve legislative facts, it is hard to justify barring appellate consideration of materials that were referenced in the district court and were subject to prior examination and dispute by all the parties.

The Concurrence also misapplies the party-presentation principal. As we recently explained, that principle generally restricts the court from raising issues *sua sponte*, but it does not preclude a court from resolving a presented issue in a way not proposed by any party. *See United States v. Cortez-Nieto*, 43 F.4th 1034, 1052 (10th Cir. 2022). For example, what if each party in a contract dispute argues that a provision of a contract is unambiguous but they disagree on what that unambiguous meaning is? It would be wholly proper for the court to decide that the provision has a third meaning or that the provision is ambiguous and further factual development is necessary.

Here, the government argues that AFTE methodology—as a whole—is reliable (so admissibility should turn only on the expertise of the examiner and the application of the methodology), while Defendant argues that the methodology is never reliable enough to be used in court. In reviewing the issue, we are not limited

to choosing between those alternatives. We can, and do, reserve ruling on the reliability of the AFTE methodology in general, largely because of the prestige of the government-sponsored reports questioning whether the accuracy of the methodology has been sufficiently tested. On the other hand, there has been considerable testing, over an extended period of time, of the accuracy of one subset of AFTE methodology—the CMS methodology used in this case. We therefore can affirm the conviction here without going further and addressing the more general question of reliability of AFTE methodology.

Given that liberty is at stake in this case, and in most cases where expert firearm toolmark testimony is offered, we are reluctant to go beyond what is necessary in this case and provide a judicial imprimatur to AFTE methodology in general. Although the concurrence believes that the majority opinion takes appellate review beyond its proper sphere, we think our approach is fully consonant with traditional appellate practice and, in the circumstances, not only prudent but less adventurous than full endorsement of the district-court ruling. We are particularly reluctant to adopt the ad hoc approach of the Concurrence, which would admit the testimony in this case while reserving judgment on admissibility in other cases, without providing needed guidance to the lower courts.

Of course, appellate judges must be particularly careful in vetting scientific or technical research that was not debated in district court. But nothing precluded Defendant, who had retained his own ballistics expert, from challenging the cited literature. In any event, it is not uncommon for an appellate court to perform an

independent *Daubert* analysis, as when it is assessing whether a district court's failure to perform the analysis was harmless error. *See Smith v. Jenkins*, 732 F.3d 51, 64–68 (1st Cir. 2013); *Sarkees v. E. I. Dupont De Nemours & Co.*, 15 F.4th 584, 589–93 (2d Cir. 2021); *UGI Sunbury LLC v. A Permanent Easement for 1.7575 Acres*, 949 F.3d 825, 833–36 (3d Cir. 2020); *Sardis v. Overhead Door Corp.*, 10 F.4th 268, 285–96 (4th Cir. 2021); *Anderson v. Raymond Corp.*, 59 F.4th 279, 283–85 (7th Cir. 2023); *United States v. Ruvalcaba-Garcia*, 923 F.3d 1183, 1190–91 (9th Cir. 2019). This court explicitly approved that practice in *StorageCraft Technology Corp. v. Kirby*, 744 F.3d 1183, 1190–92 & n.2 (10th Cir 2014) (Gorsuch, J.).

What the majority opinion does in this case is significantly less comprehensive. We have merely supplemented the district court's analysis from the *Daubert* hearing with the new information provided at trial. As previously noted, it is accepted practice for appellate courts to affirm a ruling on a pretrial motion in limine after considering what happened at trial. *See, e.g., Kinser*, 184 F.3d at 1271 (In reviewing the admissibility of expert testimony, “we do not think it necessary to confine our review to the materials accompanying the *Daubert* hearing request. Rather, we believe we may look at the entire record, including testimony presented at trial.”). At trial, expert witness Kong reported for the first time his use of the CMS methodology in comparing the cartridge shells and he described the leading study supporting that methodology. If there were flaws in that methodology or the studies, Defendant could have raised challenges; but he did not even cross-examine the expert. The concurrence complains that neither the parties nor the district court

discussed or analyzed the studies relevant to the CMS technique that were described on the AFTE website referenced by both the government and the court. But when Defendant raises no challenge to those studies or the methodology, we can properly assume their validity. Our approach, which the concurrence disapproves, has been simply to conduct our own examination of the relevant studies to confirm that the CMS methodology is not subject to the criticisms of firearm toolmark methods in general which have been made by distinguished panels of forensic experts. We see that as our duty, not a usurpation of power.

In short, we determine that the support the district court found for the general AFTE methodology is particularly strong with respect to the specific CMS methodology used in this case, making it unnecessary to resolve the reliability of the general methodology.

We recognize that our standard of review under *Daubert* is abuse of discretion. But the issues are much too important to be resolved by a simple wave of the hand and deference to the decision below. Respect for the courts is significantly reduced when litigation is resolved on the basis of pseudoscience that is rejected by highly educated and intelligent citizens who begin to think that the overlap between truth and judicial judgment is too slim. Appellate judges have an obligation to educate themselves and engage in the sometimes difficult work of assessing scientific and technical material. At least in this case, the analysis is not particularly difficult because of the extensive testing of the CMS methodology.

III. CONCLUSION

We **AFFIRM** Defendant's convictions.

United States v. Hunt, 21-6046

MORITZ, J., concurring.

I agree that the district court did not abuse its discretion when it deemed Howard Kong's toolmark-identification methodology reliable under *Daubert v. Merrell Dow Pharmaceuticals, Inc.*, 509 U.S. 579 (1993). But unlike the majority, I would reach that conclusion without supplementing the district court's analysis or engaging in an independent review of studies and literature about consecutive matching striae (CMS).

Initially, I can't join the majority's reliance on and discussion of the CMS studies because those studies are not part of the record in this case. *Cf. United States v. Kennedy*, 225 F.3d 1187, 1191 (10th Cir. 2000) ("This court will not consider material outside the record before the district court."). To be sure, citations to most of the studies discussed by the majority appear on a website that the government cited once in its response to the motion to exclude and that the district court cited once in its order. This website, published by the Association of Firearm and Tool Mark Examiners (AFTE), offers "literature citations for the more important studies that qualify as material principally concerned with the validity of firearm and toolmark identification," including "[a] short summary follow[ing] each citation." *Testability of the Scientific Principle*, The Association of Firearm and Tool Mark Examiners, <https://afte.org/resources/swggun-ark/testability-of-the-scientific-principle> (last visited March 23, 2023).

But the government's and district court's citations merely supported the *existence* of the studies listed on the website, not their quality or content. The government cited the

website to support its assertion that “[t]he AFTE theory is regularly tested both on an individual level, by peer review and verification, and on a larger level with numerous studies.” R. vol. 1, 192. Likewise, the district court cited the website to support the conclusion that “the theory of firearm toolmark identification can be and has been tested.” *United States v. Hunt*, 464 F. Supp. 3d 1252, 1257 (W.D. Okla. 2020). Thus, contrary to the majority’s characterization—that Kong’s CMS methodology is supported “by several studies referenced in the government’s pretrial *Daubert* brief and in the district court’s [o]rder”—neither the government nor the district court “referenced” any of the studies the majority relies on. Maj. Op. 27.

Nor are the government’s and district court’s passing citations to the AFTE website sufficient to incorporate any or all of the listed studies—over 100 studies grouped into six topics—into the record for purposes of appeal. The majority cites no authority to support its decision to do so, and I have found none.¹ Although a reviewing court may sometimes affirm the admission of expert testimony despite possibly flawed reasoning by the district court, such harmlessness must be “readily apparent *from the record.*” *StorageCraft Tech. Corp. v. Kirby*, 744 F.3d 1183, 1190–91 (10th Cir. 2014) (emphasis added); *see also Kinser v. Gehl Co.*, 184 F.3d 1259, 1271 (10th Cir. 1999).

¹ The majority also fails to explain whether it reviewed all the website’s listed studies or how it chose the handful that it relies on. Nevertheless, because I would not consider such studies at all, I will not substantively engage with the majority’s assessment of the literature.

Because the studies analyzed, discussed, and applied by the majority are not part of the record, I can't join the majority's decision to rely upon them.²

Moreover, and perhaps more critically, even if these studies are part of the record, neither the district court nor the parties ever mentioned such studies in their briefing below, let alone discussed or analyzed those studies. As a result, the court best positioned to evaluate the CMS literature and its reliability (the district court) had no opportunity to consider it. *See Goebel v. Denver & Rio Grande W. R.R. Co.*, 215 F.3d 1083, 1088–89 (10th Cir. 2000) (“Performance of the gatekeeping function on the record [e]nsures that a judgment in favor of either party factors in the need for reliable and relevant scientific evidence. It is not an empty exercise; appellate courts are not well-suited to exercising the discretion reserved to district courts.”). And the same is true on appeal: We lack the benefit of the parties' positions on this literature and the reliability of the CMS method. Thus, the majority's reasoning—which turns heavily, if not entirely, on its “own examination” of CMS literature that is included amongst a variety of other toolmark-identification studies on an AFTE website, Maj. Op. 41—significantly departs from “the principle of party presentation.” *United States v. Sineneng-Smith*, 140 S. Ct. 1575, 1578 (2020). This principle “restricts courts from raising new issues,” yet the majority here does exactly that when it affirms the district court's admission of Kong's expert

² The majority also appears to rely on Kong's brief mention during his trial testimony of “a study done some time ago by Biasotti and Murdock.” Maj. Op. 26 (quoting R. vol. 3, 462). But this passing mention doesn't meaningfully incorporate the majority's swath of scientific literature into the appellate record. Additionally, the study Kong mentioned doesn't appear to be included on the AFTE website, which lists one study by Biasotti and two by Murdock, but none by both individuals.

testimony based on a theory of CMS reliability that no party has either “raised” or “responded to.” *United States v. Cortez-Nieto*, 43 F.4th 1034, 1052 (10th Cir. 2022) (emphasis omitted). In so doing, the majority does far more than “supplement[] the district court’s analysis”; it ignores the district court’s analysis and conducts its own. Maj. Op. 40.

Because of these concerns, I would follow the parties’ lead and analyze the district court’s actual application of the *Daubert* factors. Doing so, I’m not persuaded that the district court erred; “nothing in the controlling legal authority” requires the “extremely high degree of intellectual purity” that Dominic Hunt presses on appeal. *United States v. Baines*, 573 F.3d 979, 989 (10th Cir. 2009). To be sure, neither this concurrence nor the majority opinion should be read to endorse the rote admission of toolmark-identification experts. *See United States v. Smith*, 756 F.3d 1179, 1193 n.13 (10th Cir. 2014) (“Firearm toolmark analysis has recently come under attack for depending on subjective judgment, rarely using control weapons, and risking an observer effect.”). But I see no abuse of discretion in the district court’s decision in this case. *See United States v. Foust*, 989 F.3d 842, 847 (10th Cir. 2021) (finding no abuse of discretion in admitting handwriting expert despite “criticism of handwriting expertise in both the courts and academic literature”). I accordingly concur.