ABRUQUAH AND THE UNCERTAIN FUTURE OF FORENSIC FEATURE-COMPARISON EVIDENCE

TABLE OF CONTENTS

INTRODUCTION

The admissibility of forensic evidence in the courtroom has been a long-standing source of contention among scientists, attorneys, and scholars.[1] Forensic science is science "used for the purposes of the law,"[2] and forensic scientists are those tasked with examining evidence to form conclusions that assist investigators in understanding what occurred during the commission of a crime.[3] A large subset of forensic science is "feature-comparison," which entails analyzing footwear impressions, hair and fibers, latent fingerprints, toolmarks, and DNA evidence.[4] The ultimate goal of forensic examiners in these disciplines is to "determine whether an evidentiary sample . . . is or is not associated with a potential source sample . . . based on the presence of similar patterns, impressions, features, or characteristics in the sample and the source."[5] Common scenarios include identifying a specific firearm as having fired certain ammunition or naming a specific person as the source of a fingerprint or DNA material.

Although they frequently seem the same, a forensic analyst's goal should not be to assist in convicting a specific suspect but rather to enhance understanding of a crime, enabling attorneys to seek justice for it.[6] This goal is derived from the ethical and scientific obligation imposed on forensic scientists to make findings supported by

1. *See generally* Jonathan J. Koehler, *How Trial Judges Should Think About Forensic Science Evidence*, 102 JUDICATURE 28, 32–33 (2018) (discussing various groups, such as the National District Attorneys Association and the American Society of Crime Laboratory Directors, and their disagreement with the findings of a disapproving report by the President's Council of Advisors on Science and Technology).

2. *What Is Forensic Science?*, AM. ACAD. FORENSIC SCIS. (2025), https://perma.cc/9M6Q-Y7LZ.

3. *Forensic Science*, U.S. DEP'T JUST. (2025), https://perma.cc/SY6H-D77B.

4. PRESIDENT'S COUNCIL OF ADVISORS ON SCI. & TECH., EXEC. OFFICE OF THE PRESIDENT, FORENSIC SCIENCE IN CRIMINAL COURTS: ENSURING SCIENTIFIC VALIDITY OF FEATURE-COMPARISON METHODS 23 (2016) [hereinafter PCAST REPORT], https://perma.cc/T8Z3-RTUZ.

5. *Id.*

6. *What Is Forensic Science?*, *supra* note 2.

objective, unbiased analysis.[7] When analysts deviate from a standard of objectivity, their credibility as an expert witness is easily challenged. Thus, the empirical validity and replicability of the analysis involved in different forensic disciplines has caused each to face varying degrees of scrutiny. For instance, DNA analysis is considered the "gold standard" of forensic science and faces few courtroom admissibility issues because of its reliable and objective methodology.[8] In contrast, bitemark evidence, while once accepted, is now considered unreliable and scientifically invalid since its analysis requires a large degree of subjectivity.[9]

Firearm identification, one of the most longstanding and well-established (albeit controversial) branches of forensic science, has faced increasingly vehement challenges to its admissibility in court over the past two decades. Following strong critiques in official reviews of the discipline in 2009 and 2016,[10] legal scholars have routinely called for judges to renounce the current almost-guaranteed admissibility of firearm evidence in courtrooms.[11] Yet, it has only been within the last few years that courts have begun to heed this advice and reconsider the evidence's validity and reliability under the *Daubert* standard.[12]

This trend culminated in the Maryland Supreme Court's 2023 decision in *Abruquah v. State*,[13] where it applied the *Daubert-Rochkind* factors to firearm evidence.[14] It held that the reliability of

---

7. *Id.*

8. *See* PCAST REPORT, *supra* note 4, at 26; Celia Henry Arnaud, *Thirty Years of DNA Forensics: How DNA Has Revolutionized Criminal Investigations*, C&EN (Sept. 18, 2017), https://perma.cc/YH9F-EJ8M.

9. *See* PCAST REPORT, *supra* note 4, at 87; Daniele Selby, *Why Bite Mark Evidence Should Never Be Used in Criminal Trials*, INNOCENCE PROJECT (Apr. 26, 2020), https://perma.cc/DP6L-3K3D.

10. *See generally* COMM. ON IDENTIFYING THE NEEDS OF THE FORENSIC SCI. CMTY., NAT'L RSCH. COUNCIL OF THE NAT'L ACADS., STRENGTHENING FORENSIC SCIENCE IN THE UNITED STATES: A PATH FORWARD (2009) [hereinafter NAS REPORT], https://perma.cc/KL5C-LPLF; PCAST REPORT, *supra* note 4.

11. *See generally* Jennifer Friedman & Jessica Brand, *It Is Now up to the Courts: "Forensic Science in Criminal Courts: Ensuring Scientific Validity of Feature-Comparison Methods,"* 57 SANTA CLARA L. REV. 367 (2017); Brandon L. Garrett et al., *Judging Firearms Evidence*, 97 S. CAL. L. REV. 101 (2024); Bonnie Lanigan, Note, *Firearms Identification: The Need for a Critical Approach To, and Possible Guidelines For, the Admissibility of "Ballistics" Evidence*, 17 SUFFOLK J. TRIAL & APP. ADVOC. 54 (2012). To be sure, challenges to the admissibility of firearm evidence existed prior to the publication of the NAS and PCAST reports. However, as the articles in this footnote discuss, the challenges never had the teeth necessary to support them until these reports showcased to everyone the major shortcomings of the discipline.

12. Daubert v. Merrell Dow Pharms., Inc., 509 U.S. 579 (1993); *see infra* Part III.

13. 296 A.3d 961 (Md. 2023).

14. *Id.* at 971.

the methodology underlying "identification" conclusions that a specific firearm fired "unknown" bullets was not scientifically supported.[15] The holding created the first state-wide ban on firearm identification testimony relating to bullets. This seemingly new scrutiny and more rigid application of *Daubert* to firearm evidence casts doubt on the future admissibility of feature-comparison evidence.

Below, Part I details forensic firearm examination, and Part II overviews the historical acceptance of forensic science in the law. Part III then explores the shift in permissible forensic conclusions, the resulting outright exclusion of certain firearm evidence, and the *Abruquah* court's majority and dissenting opinions. Next, Part IV addresses the validity of the court's analyses and agrees that the Majority properly excluded the identification testimony. Finally, in Part V, the ramifications of the *Abruquah* decision are discussed. First, it reports how courts in Maryland have interpreted *Abruquah* in subsequent cases involving similar evidentiary challenges. Then, it considers the broader implications of the Majority's reasoning and holding, including how additional courts may begin to similarly exclude not only firearm evidence but also other forensic feature-comparison evidence. Ultimately, this Note concludes that, where a defendant's liberty is at stake, the cost of allowing unsubstantiated firearm "identification" conclusions is too great. Therefore, until a comprehensive study of mass-produced firearm similarities facilitates a more detailed methodology to guide analysts, this testimony cannot be admissible.

I. THE DEVELOPMENT OF FIREARM FEATURE-COMPARISON EVIDENCE

The underlying premise behind firearm identification is that each gun leaves toolmarks on the ammunition components it fires, which are unique compared to all other firearms.[16] A toolmark is an impression left in a softer material when force is applied to it by a harder object, or tool.[17] The following Sections detail the impressions created when a firearm is discharged and how those markings are compared in casework.

15. *Id.* at 968.

16. Jinsong Leng & Zhihu Huang, *On Analysis of Circle Moments and Texture Features for Cartridge Images Recognition*, 39 EXPERT SYS. WITH APPLICATIONS 2092, 2092 (2012).

17. NAS REPORT, *supra* note 10, at 150; ROBERT M. THOMPSON, FIREARM IDENTIFICATION IN THE FORENSIC SCIENCE LABORATORY 15 (2010).

A. What Is Being Compared?

When a firearm's trigger is pulled, the hammer strikes the firing pin, slamming it into the primer at the base of the cartridge case.[18] This strike ignites the powder inside the ammunition, causing an accumulation of gas pressure.[19] The resulting force thrusts the bullet down the barrel while simultaneously pushing the cartridge case back against the breech face.[20] Then, the cartridge case is ordinarily removed from the chamber by an extractor and subsequently expelled from the firearm by an ejector.[21]

During each of these processes, impressions are created on the ammunition components. For instance, the inside of a barrel has spiral grooves that help stabilize the bullet as it is expelled.[22] Consequently, as a bullet travels the length of the barrel, these grooves become impressed into it, creating rifling marks.[23] Additionally, the firing pin, breech face, and ejector each produce impressions on the base of the cartridge case.[24] Then, when the extractor hook latches onto the rim of the cartridge case, a grip impression is created along the cartridge case's side.[25] Therefore, the firing of a weapon will generally result in four distinct impressions on a cartridge case as well as striations on a bullet.[26]

B. How Are They Compared?

The Association of Firearm and Tool Mark Examiners (AFTE) has produced the most commonly used methodology for firearm examination, the AFTE Theory of Identification, which requires microscopic comparison of the impressions on "unknown" evidentiary items to "known" source items.[27] The Theory can be distilled into two overall steps: (1) evaluation of class characteristics to exclude non-sources and (2) comparison of individual and subclass characteristics to identify a specific source firearm.[28]

---

18. THE ROYAL SOC'Y, UNDERSTANDING BALLISTICS: A PRIMER FOR COURTS 22 (2021), https://perma.cc/NG7J-NCXL.

19. *Id.*

20. *Id.*

21. *Firearms Examiner Training: Glossary*, NAT'L INST. JUST. (2025), https://perma.cc/JE2H-9MZ8.

22. NAS REPORT, *supra* note 10, at 151.

23. *Id.*

24. *Id.*

25. *Forensic Marks on a Cartridge Case*, NIST (Nov. 1, 2021), https://perma.cc/2X3M-D9NT.

26. *See* THE ROYAL SOC'Y, *supra* note 18, at 30. For an excellent visualization of the mechanics of shooting a firearm, as well as the impressions that result on a cartridge case, see *Forensic Marks on a Cartridge Case*, *supra* note 25.

27. Abruquah v. State, 296 A.3d 961, 974 (Md. 2023); PCAST REPORT, *supra* note 4, at 104.

28. PCAST REPORT, *supra* note 4, at 104.

Class characteristics are traits held by many items of the same type.[29] With respect to firearms, class characteristics include the shape of a firing pin as well as the number, width, and twist of grooves in a barrel.[30] These traits are common to all firearms of the same make and model because they are features the manufacturer selects when designing a specific firearm.[31] Thus, while merely examining the class characteristics observed on a spent cartridge case or bullet would not allow a forensic examiner to identify a source firearm, they are sufficient to exclude a known firearm as a possible source.[32]

If an exclusion cannot be made, examiners must then identify and compare the subclass and individual characteristics observed on the samples.[33] Subclass characteristics are markings produced in the manufacturing process that are similar among a small group of firearms, such as impressions created by a dull tool.[34] However, they are not features specifically chosen by the manufacturer like class characteristics.[35]

In contrast to both class and subclass characteristics, individual characteristics are those microscopic markings that are unique to a specific firearm.[36] These irregularities and imperfections can result from the production process or the normal wear-and-tear of the firearm.[37] Although modern production processes have greatly increased the uniformity between mass-produced items, in theory there are still microscopic differences between items because of the gradual erosion of the manufacturing equipment after each use.[38] Under this assumption, even sequentially produced firearms would have slight differences in their internal components which would, in turn, transfer onto the bullet and cartridge case during the firing process.

The AFTE Theory calls for the use of a comparison microscope to simultaneously view the markings on a "known" and "unknown" bullet or cartridge case to compare them.[39] Under the Theory, a firearm examiner can reach one of four conclusions: identification,

29. NAS REPORT, *supra* note 10, at 152.

30. *Id.*

31. PCAST REPORT, *supra* note 4, at 104; ASS'N OF FIREARM & TOOL MARK EXAM'RS, GLOSSARY 39 (6th ed. 2013), https://perma.cc/H56Z-5GRY.

32. PCAST REPORT, *supra* note 4, at 104; *AFTE Range of Conclusions*, ASS'N FIREARM & TOOL MARK EXAM'RS (2025), https://perma.cc/FR7M-GNZ7.

33. PCAST REPORT, *supra* note 4, at 104.

34. NAS REPORT, *supra* note 10, at 152.

35. ASS'N OF FIREARM & TOOL MARK EXAM'RS, *supra* note 31, at 122.

36. NAS REPORT, *supra* note 10, at 152.

37. ASS'N OF FIREARM & TOOL MARK EXAM'RS, *supra* note 31, at 66.

38. NAS REPORT, *supra* note 10, at 150.

39. *Id.* at 152. A comparison microscope is "made from two compound microscopes joined by a comparison bridge that allows viewing of two objects at the same time." *Id.*

inconclusive, elimination, or unsuitable.[40] Relevant to this Note, an examiner may make an "identification" (a conclusion that one weapon discharged a specific bullet or cartridge case) when there is "sufficient agreement" between the impressions on the "known" and "unknown" samples.[41] Sufficient agreement exists when "the agreement of individual characteristics is of a quantity and quality that the likelihood another tool could have made the mark is so remote as to be considered a practical impossibility."[42]

However, the Theory does not provide guidelines for examiners to follow when classifying markings as individual or subclass characteristics. Moreover, the Theory does not detail what satisfies the "quantity and quality" of agreement required for identification. Instead, it states that significant agreement exists "when the agreement in individual characteristics exceeds the best agreement demonstrated between toolmarks known to have been produced by different tools and is consistent with agreement demonstrated by toolmarks known to have been produced by the same tool."[43] Indeed, it explicitly acknowledges that it employs a subjective process reliant upon examiners' "training and experience."[44] Thus, the Theory's entire methodology for forming "identification" conclusions relies on an examiner's subjective judgments at each step of the analysis.

II. THE HISTORICAL ACCEPTANCE OF FORENSIC SCIENCE IN COURTROOMS

The forensic discipline of firearm identification has existed for over a century.[45] In the early 1900s, courts diverged on whether to admit firearm identification evidence.[46] However, beginning in the late 1920s and early 1930s, forensic scientist Calvin Goddard developed new methods for firearm examination using a comparison microscope.[47] He testified that examiners using his methodology could reliably conclude whether a firearm had fired "unknown" spent

40. *AFTE Range of Conclusions*, *supra* note 32. For the definition of each accepted conclusion, see *Firearms Examiner Training: AFTE Theory of Identification*, NAT'L INST. JUST. (2025), https://perma.cc/7FX7-H2GP.

41. *Firearms Examiner Training: AFTE Theory of Identification*, *supra* note 40; *Theory of Identification as It Relates to Toolmarks*, ASS'N FIREARM & TOOL MARK EXAM'RS (Sept. 16, 2024), https://perma.cc/R4FN-54RV.

42. *Theory of Identification as It Relates to Toolmarks*, *supra* note 41.

43. *Id.*

44. *Id.*

45. Abruquah v. State, 296 A.3d 961, 976 (Md. 2023); *see* Garrett et al., *supra* note 11, at 110.

46. Garrett et al., *supra* note 11, at 113–14.

47. *Id.* at 119–20. Calvin Goddard is now referred to as the "father of forensic ballistics." *Then: Law School Lab Advanced Study of Ballistics*, NW. MAG.: CAMPUS LIFE (2016), https://perma.cc/U4SH-YXWU.

ammunition.[48] Since then, courts have been largely accepting of firearm identification evidence, allowing examiners to testify as expert witnesses without much inquiry into their qualifications.[49]

It was not until the 1990s, when new DNA analysis technologies emerged, that there began "serious questioning of the validity of many of the traditional forensic disciplines."[50] Around the same time, the Supreme Court decided *Daubert v. Merrell Dow Pharmaceuticals, Inc.*,[51] which provided five considerations for judges to assess the reliability, and thus admissibility, of scientific evidence: (1) whether the technique can be, and has been, tested; (2) whether the technique has been peer reviewed and published; (3) the technique's known or potential error rate; (4) the existence and maintenance of controls and standards; and (5) whether the technique has garnered widespread acceptance within the relevant scientific community.[52]

Then, in 2009, the National Academy of Sciences (NAS) published a comprehensive review of forensic science disciplines, with observations about how they met—or more commonly, failed to meet—the *Daubert* factors.[53] With regard to firearm identification, the NAS Report reiterated the findings of its 2008 study that the "validity of the fundamental assumptions of uniqueness and reproducibility of firearms-related toolmarks has not yet been fully demonstrated."[54] More specifically, the NAS Report stated that "[s]ufficient studies have not been done to understand the reliability and repeatability of the methods," and suggested that these studies were necessary to enable accurate testimony about individualization and source identification.[55] Further, the NAS Report stated that the AFTE Theory lacked "a specific protocol" and criticized it for not even addressing "questions regarding variability, reliability, repeatability, or the number of correlations needed" to instill trust in conclusions.[56]

---

48. Garrett et al., *supra* note 11, at 110–11.

49. *Id.* at 113–20; *Abruquah*, 296 A.3d at 976. Aptly stated, "[i]f the State could produce a witness in the mold of Major Goddard, following the now-respected comparison microscope methodology, then the testimony would routinely be admitted." Garrett et al., *supra* note 11, at 120.

50. PCAST REPORT, *supra* note 4, at 25.

51. 509 U.S. 579 (1993).

52. *Id.* at 593–94.

53. *See generally* NAS REPORT, *supra* note 10.

54. *Id.* at 154 (quoting NAT'L RSCH. COUNCIL OF THE NAT'L ACADS., BALLISTIC IMAGING 81 (Daniel L. Cork et al. eds., 2008)). Reproducibility refers to "the likelihood that, for a particular sample, a different examiner will make the same determination as the initial examiner." *Abruquah*, 296 A.3d at 1021 (Gould, J., dissenting).

55. NAS REPORT, *supra* note 10, at 154. Repeatability refers to "the likelihood that the same examiner will make the same determination for a particular sample on a subsequent examination." *Abruquah*, 296 A.3d at 1017 (Gould, J., dissenting).

56. NAS REPORT, *supra* note 10, at 155.

Following the NAS Report's scathing criticism, forensic scientists should have been motivated to restore trust in their fields by addressing the identified weaknesses. Accordingly, in 2016, the President's Council of Advisors on Science and Technology (PCAST) set out to reevaluate the problematic feature-comparison disciplines in light of any new research.[57] However, disappointingly, the PCAST Report showcased an overall lack of advancement in the field of firearm identification, with no significant studies having been conducted to further support the discipline's reliability or validity.[58]

In fact, PCAST noted that existing studies "seriously underestimate[d]" false-positive rates and that only one study, the Ames I Study, had been "appropriately designed."[59] Significantly, since that study had not yet been subject to peer review or publication, its reproducibility was unknown and it could not constitute a sufficient basis for foundational validity.[60] Ultimately, because of this, the Report called for courts to recognize that the high accuracy rate touted by firearm examiners was not an authentic representation of the likelihood of errors in casework.[61]

Furthermore, the PCAST Report strongly criticized the AFTE's Theory as being "circular."[62] It stated that "[t]he frequency with which a particular pattern or set of features will be observed in different samples, which is an essential element in drawing conclusions, is not a matter of 'judgment.'"[63] It continued that "an expert's expression of *confidence* based on personal professional experience or expressions

---

57. *See* PCAST REPORT, *supra* note 4, at 22–23.

58. *See id.* at 111.

59. *Id.*; *see Abruquah*, 296 A.3d at 978. In this context, "appropriately designed" refers to black box, or open-set, studies, which are those in which "there is no guarantee that the correct source is present." PCAST REPORT, *supra* note 4, at 108. In contrast, closed-set studies are those in which "the correct source is present for each questioned sample[]." *Id.* at 86. Because open-set studies are most like the casework that forensic examiners regularly perform, they are the preferable study design and are thought to produce more realistic false-positive rates. *Id.* at 108. When the PCAST Report came out in 2016, the Ames I Study was the only open-set study, and thus the only "appropriately designed" study, in the field of firearm identification. *Id.* at 111.

60. PCAST REPORT, *supra* note 4, at 111. Foundational validity means that a study is (1) reproducible, (2) repeatable, and (3) accurate. *Id.* at 47. The PCAST Report stated that, because "[t]he scientific criteria for foundational validity require appropriately designed studies by *more than one group* to ensure reproducibility[,] . . . additional, appropriately designed black-box studies" were necessary to support the discipline. *Id.* at 111. In response to PCAST's commentary, the Ames II Study, a second open-set, black box firearm identification study, was conducted in 2020. *Abruquah*, 296 A.3d at 980. The Study "was designed specifically to test the repeatability and reproducibility of the AFTE Theory methodology." *Id.* at 993.

61. *See* PCAST REPORT, *supra* note 4, at 112.

62. *Id.* at 104.

63. *Id.* at 55.

of *consensus* among practitioners about the accuracy of their field is no substitute for error rates estimated from relevant studies."[64] The PCAST Report concluded that "establishing foundational validity based on empirical evidence is thus a *sine qua non*. Nothing can substitute for it."[65]

III. CHANGING CONCLUSIONS IN FORENSIC SCIENCE: FROM "MATCH" TO PROBABILITY

Following the findings in the NAS and PCAST reports, defense attorneys became more tenacious and forceful in their objections to the admission of forensic feature-comparison evidence in court, arguing that the disciplines failed to meet evidentiary standards.[66] Still, only recently have judges begun to truly consider *Daubert* and Federal Rule of Evidence (FRE) 702 when determining the admissibility of firearm evidence in their courtrooms.[67]

*A. Shifting Language Pre-*Abruquah*: 2005–2022*

While almost all judges have hesitated to outright exclude forensic evidence from their courtrooms,[68] Professor Brandon Garrett identified thirty-seven cases between 2005 and 2022 where limitations were imposed on forensic firearm examiners' testimony.[69] The restrictions enforced by courts are generally intended either to narrow the language an expert may use to phrase their conclusion or only allow conclusions regarding class characteristics.[70]

For example, instead of allowing "match" testimony or conclusions that denote a degree of certainty that a firearm fired the "unknown" bullet or cartridge case, judges are now more likely to limit testimony to much less definitive terms.[71] Commonly, the

64. *Id.*

65. *Id.* at 6.

66. Garrett et al., *supra* note 11, at 141.

67. *Id.* at 160.

68. The first, and only, court to completely exclude firearm identification evidence just recently had its landmark decision vacated. People v. Winfield, No. 15-CR-14066-01, slip op. at 36 (Ill. Cir. Ct. Feb. 8, 2023), https://perma.cc/3SX2-E3TH (finding "no objective forensic based reasons that firearms identification evidence belongs in any category of forensic science" and limiting the use of firearm evidence to police investigations outside the courtroom), *vacated*, slip op. at 26 (Dec. 18, 2024), https://perma.cc/M6R2-K4KF (concluding that "[f]irearms identification evidence is generally accepted, admissible and relevant" and only prohibiting experts from testifying to "a reasonable degree of scientific certainty"). *See* Radley Balko, *A Chicago Judge Just Erased Her Predecessor's Historic Ruling on Forensic Firearms Analysis*, WATCH (Jan. 14, 2025), https://perma.cc/5YW8-WSKL.

69. Garrett et al., *supra* note 11, at 141–42, 147–48.

70. *Id.* at 147.

71. *See, e.g.*, United States v. Adams, 444 F. Supp. 3d 1248, 1266–67 (D. Or. 2020) (prohibiting testimony of a match after finding that the conclusion was not

limitations are to conclusions of "more likely than not"[72] or "consistent with."[73] This trend away from allowing definitive statements of source identification towards only permitting general, probabilistic conclusions is also reflected in the Department of Justice's recent disallowance of using certainty-based language in conclusions: "An examiner shall not . . . use the expressions 'reasonable degree of scientific certainty,' 'reasonable scientific certainty,' or similar assertions . . . ."[74] This trend, though, is not unique to the discipline of firearm identification—for example, the language has also shifted for DNA analysis, which no longer allows "identification" conclusions, regardless of the likelihood ratio produced.[75]

Additionally, a handful of courts have proscribed firearm examiners from testifying to anything more than class characteristics due to the AFTE Theory's lack of formal procedure for comparing individualizing characteristics.[76] This means that examiners may merely testify that "the same *type* of gun fired the bullets or cartridge cases, but . . . cannot say that the *same* gun fired the bullets or

---

the result of a "*scientific* inquiry"); Williams v. United States, 210 A.3d 734, 740–43 (D.C. 2019) (finding error in allowing an examiner to testify to the absolute certainty that a firearm was the source to the exclusion of all other firearms). *Compare* United States v. Shipp, 422 F. Supp. 3d 762, 783 (E.D.N.Y. 2019) (disallowing testimony "to any degree of certainty"), *with* United States v. Gil, 680 F. App'x 11, 14 (2d Cir. 2017) (permitting testimony to a "reasonable degree of certainty" after recognizing that the discipline "could not fairly be referred to as 'scientific'" (citation omitted)), *and* United States v. Ashburn, 88 F. Supp. 3d 239, 250 (E.D.N.Y. 2015) (allowing conclusions within a "reasonable degree of certainty").

72. *See, e.g.*, Whitfield v. Riley, No. 09-1877, 2021 WL 1795554, at *4 (E.D. La. May 5, 2021); United States v. Glynn, 578 F. Supp. 2d 567, 575 (S.D.N.Y. 2008).

73. *See, e.g.*, United States v. Davis, No. 18-cr-00011, 2019 WL 4306971, at *6–7 (W.D. Va. Sept. 11, 2019); *see also* Garrett et al., *supra* note 11, at 148–49, 162 (listing additional cases where conclusions were limited to "consistent with" language). *But see* United States v. Briscoe, 703 F. Supp. 3d 1288, 1308 (D.N.M. 2023) (prohibiting both "match" and "consistent with" language and only allowing testimony that "areas of comparison . . . are similar").

74. U.S. DEP'T OF JUST., UNITED STATES DEPARTMENT OF JUSTICE UNIFORM LANGUAGE FOR TESTIMONY AND REPORTS FOR THE FORENSIC FIREARMS/TOOLMARKS DISCIPLINE PATTERN EXAMINATION 3 (2023), https://www.justice.gov/olp/media/1295781/dl.

75. *See* U.S. DEP'T OF JUST., DEPARTMENT OF JUSTICE UNIFORM LANGUAGE FOR TESTIMONY AND REPORTS FOR FORENSIC AUTOSOMAL DNA EXAMINATIONS USING PROBABILISTIC GENOTYPING SYSTEMS 4 (2022), https://www.justice.gov/media/969526/dl.

76. *See, e.g.*, *Adams*, 444 F. Supp. 3d at 1267 (stating that firearm examiners could only testify to the caliber, shape of the firing pin, number of lands and grooves, and direction of twist observed on the firearm and the "known" and "unknown" samples).

cartridge cases."[77] Limiting the testimony in this way necessitates an objective conclusion that does not rely upon what many people believe to be unfounded scientific principles, such as attempting to identify and compare individual characteristics.

B. The Abruquah *Majority's Opinion*

Recently, the Maryland Supreme Court became the first state high court to ban the admittance of firearm identification testimony relating to bullets in criminal matters.[78] In *Abruquah*, a first-degree murder case stemming from 2013, the defendant moved to exclude expert testimony that his gun fired the bullet fragments found at the crime scene.[79] After the trial court denied the motion and he was convicted, Mr. Abruquah appealed.[80] As his appeal was pending, the Maryland Supreme Court decided *Rochkind v. Stevenson*,[81] which "abandoned the *Frye-Reed* standard for admissibility . . . in favor of the [*Daubert*] standard."[82] Therefore, the case was remanded for consideration of whether the evidence would have been admissible under the new ten-factor *Daubert-Rochkind* standard, which consisted of the five *Daubert* factors and the five factors listed in the Advisory Committee Notes to the 2000 amendment of FRE 702.[83]

After the lower court affirmed the evidence's admission, the Maryland Supreme Court granted cert to determine "whether the AFTE Theory [could] reliably support an unqualified opinion that a particular firearm is the source of one or more particular bullets" under the new evidentiary standard.[84] In June 2023, the court held that a firearm examiner cannot testify to a specific firearm as being the weapon that discharged an "unknown" evidentiary bullet.[85] Instead, it followed the trend of some trial and intermediate courts in only allowing an examiner to testify that the "unknown" sample is similar to the "known," without making any conclusion as to the source of the "unknown" evidence.[86]

---

77. Garrett et al., *supra* note 11, at 151.

78. *See* Abruquah v. State, 296 A.3d 961 (Md. 2023).

79. *Id.* at 969.

80. *Id.* at 970.

81. 236 A.3d 630 (Md. 2020).

82. *Abruquah*, 296 A.3d at 970. The *Frye-Reed* standard had required merely that the methodology relied upon be generally accepted by the scientific community. *Id.* at 969; Frye v. United States, 293 F. 1013, 1014 (D.C. Cir. 1923); *see The Frye "General Acceptance" Standard*, NAT'L INST. JUST. (2025), https://perma.cc/PR8R-WHMV.

83. *See Rochkind*, 236 A.3d at 650. This ten-part test is abnormal in that it emphasizes that state courts should consider FRE 702 in determining evidence reliability and apparently provides even greater flexibility in making admissibility decisions. *See Abruquah*, 296 A.3d at 999 (Hotten, J., dissenting).

84. *Abruquah*, 296 A.3d at 987 (majority opinion).

85. *Id.* at 998.

86. *Id.*

The *Abruquah* Majority concluded that the six *Daubert-Rochkind* factors weighing in favor of exclusion (factors two, three, four, seven, eight, and ten) were stronger than the neutral factor (factor five) or the three that weighed in favor of admission (factors one, six, and nine).[87] However, it repeatedly emphasized that its decision was made based on the evidence presented at the trial court and that future courts may hold differently if presented with additional studies that "materially alter[] the analysis applicable to the reliability of the [AFTE Theory]."[88]

1. Testability

First, the court concluded that the testability of firearm identification weighed in favor of admitting the evidence.[89] The fact that there have been many studies on firearm analysis demonstrated that the discipline was testable—the issue was whether the testing was appropriately designed to show reliability or reproducibility.[90]

2. Peer Review and Publication

Second, the court held that the peer review and publication factor weighed in favor of exclusion.[91] It stated that "[t]he two most robust studies . . . have not been peer reviewed or published in a journal."[92] Moreover, there was no "information about the extent or quality of peer review as concerns the validity of the methodology" of any of the other studies cited by the prosecution.[93] The court called the NAS and PCAST Reports a form of peer review but stated that their scathing critiques demonstrated the discipline's flaws.[94]

3. Known or Potential Error Rate

Third, the known or potential error rate of the discipline weighed in favor of exclusion after extensive analysis by the court.[95] This factor entails the accuracy as well as the reproducibility and reliability of results.[96] At the *Frye-Reed* hearing, the government had argued that the extremely low false-positive rates reported in the

87. *Id.* at 988–97. In contrast, the circuit court had held that all factors except four and seven weighed in favor of admission. *Id.* at 1000 (Hotten, J., dissenting).

88. *Id.* at 972 n.6 (majority opinion). In fact, the court specified that it did not consider studies the government attempted to present to the circuit court. *Id.*

89. *Id.* at 988.

90. *See id.*

91. *See id.*

92. *Id.* The court was referring to the Ames I and II Studies. *See supra* notes 59–60.

93. *Abruquah*, 296 A.3d at 988.

94. *Id.*

95. *See id.* at 989–91.

96. *See id.*

Ames I and II Studies supported the validity of the closed-set studies criticized by PCAST.[97] Indeed, the court noted that the low false-positive rates were the strongest evidence in favor of admission.[98]

However, the court ultimately agreed with the defendant that the single-digit error rates reported in the Ames Studies were unreliable.[99] The court recognized that the Ames Studies produced high rates of "inconclusive" responses, which were drastically increased from the number of reported "inconclusive" responses in the PCAST-critiqued closed-set studies.[100] Because the examiners in the study were more incentivized to not be incorrect than they would have been in normal casework (i.e., respond "inconclusive" instead of erroneously choose "identification" or "exclusion"), the court doubted the veracity of the error rates.[101] Siding with a growing number of scholars,[102] the court concluded that at least some of the "inconclusive" results should have been treated as erroneous conclusions.[103] But categorizing even one type of "inconclusive" response as erroneous would have skyrocketed the false-positive rate from 0.7% to 10.13%.[104]

The court also expressed concern at the fact that there was no indication that the low false-positive error rates were similar to the actual false-positive rates in casework.[105] This was particularly worrisome given the extreme consequences that can result from a firearm examiner testifying that a "bullet found in a victim's body was fired from the defendant's gun."[106] Moreover, the court noted that the accuracy rate between the Ruger and Beretta handguns used in the Ames II Study were wildly different, and the lack of comprehensive data for other brands of firearms made it difficult to determine what

97. *Id.* at 983.

98. *Id.* at 991.

99. *Id.* at 990.

100. See *supra* note 59 for a description of closed-set studies and a discussion of how the Ames Studies differed.

101. *Abruquah*, 296 A.3d at 990.

102. *See, e.g.*, Richard E. Gutierrez, *Bowling with Bumper Rails: How Firearms Examiners Have Duped the Courts and Generated Low Error Rates Only by Avoiding Challenging Comparisons*, 75 U.C. L.J. 1535, 1551–54 (2024).

103. *Abruquah*, 296 A.3d at 990.

104. *Id.* Part of the Majority's reasoning was that there was a ground truth in the Study—all the samples had either a match or no match, and participants knew that. *Id.* at 980, 990; *see also id.* at 1011 (Gould, J., dissenting). As such, the Majority agreed with the defendant's argument that giving a sample an "inconclusive" result when there was a correct answer should have been counted as an error. *See id.* at 990 (majority opinion); *see also id.* at 1011 (Gould, J., dissenting).

105. *Id.* at 990–91 (majority opinion).

106. *Id.* at 991.

the error rate would have been for the Taurus revolver, the firearm in question in the case.[107]

4. Existence and Maintenance of Standards and Controls

Fourth, the court held that both the lack of standards and controls as well as the minimal value of those existing favored exclusion.[108] The court discounted third-party reviewers' conclusion verifications since they were not conducted blind.[109] Further, despite the circular AFTE Theory lacking "any guiding standard other than the examiner's own subjective judgment," no evidence had been presented regarding the efficacy of proficiency testing to demonstrate that reviewers' judgment would be credible.[110]

Again, the court referenced the Ames II Study's findings on the lack of reproducibility, which it said highlighted "the absence of any standards or controls to guide the analysis of examiners."[111] The court also emphasized the lack of standards or controls guiding the classification of individual or class characteristics.[112]

5. General Acceptance

Fifth, the general acceptance of firearm identification was deemed neutral, neither weighing in favor of admission nor exclusion.[113] While firearm examiners and law enforcement are largely accepting of the AFTE Theory, the court recognized that other scholars and scientists have been extremely critical of both the validity of the discipline and its existing methodology.[114] Ultimately, it held the factor neutral since it "would be remiss . . . to rely exclusively on a community that, by definition, is dependent for its livelihood on the continued viability of a methodology to sustain it, while ignoring the relevant and persuasive input of . . . well-qualified, and disinterested" professionals.[115]

6. Whether Opinions Were Developed for Litigation

Sixth, the court concluded that the expert opinion at issue was developed for litigation, but still held that the factor supported admission.[116] It stated that the factor's purpose was to "determine whether there is reason for skepticism that the opinion reached might be tailored to the preferred result . . . rather than the expert's

107. *Id.* at 990.
108. *Id.* at 991.
109. *Id.*
110. *Id.* at 992.
111. *Id.* at 993.
112. *Id.*
113. *Id.* at 994.
114. *Id.*
115. *Id.*
116. *Id.* at 994–95.

considered, independent conclusion," and found no reason to question the examiner's integrity.[117]

7. Unjustified Extrapolation from an Accepted Premise

Seventh, the court reiterated that the case demonstrated an analytical gap between the expert's opinion and the "empirical foundation on which the opinion was derived."[118] It stated that the "unqualified statement that the bullets were fired from [the defendant's] revolver [was] still more definitive than can be supported by the record."[119] The court accepted that the AFTE Theory allowed reliable opinions of consistency between "unknown" and "known" bullets, but concluded that it did not support a reliable opinion that bullets were fired from a specific firearm.[120]

8. Accounting for Obvious Alternative Explanations

Eighth, the court held that the firearm examiner's failure to address alternative explanations supported exclusion.[121] The examiner's lack of analysis and comparison of "other firearms in the same production run as the firearm under examination" meant that he could not appropriately or reliably either (1) classify specific markings as class, subclass, or individualizing characteristics or (2) determine that other firearms would not leave those impressions.[122]

9. Level of Care

Ninth, the court held that the firearm examiner's level of care, while not applicable since he performed the comparison as part of his profession, nonetheless weighed in favor of admission.[123] The court had "no qualms about accepting . . . that [the examiner] is a 'consummate professional in his field' and demonstrated a 'level of care in this case,'" tipping the scales in favor of the government.[124]

10. Relationship Between the Reliability of Methodology and Opinion

Lastly, the minimal relationship between the reliability of the AFTE methodology and the examiner's opinion also favored exclusion.[125] Although the court acknowledged that firearms analysis is "generally reliable" in determining the consistency between

117. *Id.* at 995.
118. *Id.* (quoting State v. Matthews, 277 A.3d 991, 1014 (Md. 2022)).
119. *Id.*
120. *Id.* at 996.
121. *Id.*
122. *Id.*
123. *Id.*
124. *Id.*
125. *Id.* at 996–97.

"unknown" and "known" bullet or cartridge case evidence, it opined that the reliability of identifying a "particular known firearm" as the source of a "particular unknown bullet" was not established.[126] Moreover, it implied that firearm examiners may, in the future, be able to testify to the level of consistency between samples, only after "additional studies and data" become available.[127]

C. Justice Gould's Abruquah *Dissent*

In his nearly thirty-page dissent, Justice Gould began by strongly criticizing the Majority's treatment of the PCAST Report as "gospel" and its focus on only the Ames I and II Studies.[128] He stated that the closed-set studies should not have been written off as unreliable or unhelpful simply because the examiners knew there was a correct match when making comparisons.[129] However, for the remainder of his opinion, Justice Gould humored the Majority's focus on PCAST and argued that the Ames I and II Studies still provided the reliability that PCAST demanded.

Justice Gould expressed that the he did not see the rationale for treating "inconclusive" responses as erroneous since there was no evidence that examiners in studies are more inclined to conclude "inconclusive" simply to avoid being incorrect.[130] Instead, he advocated for a manner of calculating error rates based only on conclusive determinations, without considering "inconclusive" responses.[131] He argued that this approach aligned with the ultimate issue before the court: whether the AFTE Theory supports a conclusion as correct when an examiner determines that a specific firearm fired the "unknown" bullet.[132] In fact, the PCAST Report itself had stated that false-positive rates could be determined either "based on the *conclusive* examinations or on *all* examinations" before it ultimately recommended that "false positive rates should be based only on conclusive examinations."[133] Further, Justice Gould argued

126. *Id.*

127. *Id.* at 996.

128. *Id.* at 1007–10 (Gould, J., dissenting).

129. *Id.* at 1009–10.

130. *Id.* at 1012. Justice Gould contrasted the Ames I and II Studies with proficiency testing to demonstrate that the examiners participating in the Studies did not have any personal or professional stake in reporting correct answers. *Id.* Additionally, he noted that the examiners participating in the Studies were not told how the "inconclusive" responses would be treated, so they could not have been more inclined to denote them one way versus another. *Id.* at 1012–13.

131. *Id.* at 1013.

132. *Id.*

133. *Id.* at 1014 (quoting PCAST REPORT, *supra* note 4, at 153). The justification for this approach is that generally only conclusive evidence would be used against a defendant in court. *Id.*

that completely excluding "inconclusive" responses penalizes examiners by reducing sample sizes and increasing error rates.[134]

Relying on PCAST's framework, Justice Gould remarked that a false-positive rate of under 5.0% was generally sufficient to be reliable.[135] After analyzing both the Ames I and II Studies' reported false-positive and false-negative rates, and recalculating the rates of each to exclude "inconclusive" responses, Justice Gould noted that all were below the 5.0% threshold.[136] Thus, he concluded that the error rates in both Studies supported "that the trial court's determination of reliability was reasonable."[137]

Justice Gould then addressed the Majority's concern about the disparity between the repeatability and reproducibility rates in the Ames II Study.[138] In the Study, the repeatability rates were 79.0% (matching) and 64.7% (non-matching), while the reproducibility rates were 68.0% (matching) and 31.0% (non-matching).[139] However, Justice Gould demonstrated that the pooling of the three types of "inconclusive" conclusions was an important influence on those rates. When the "inconclusive" responses were combined, the repeatability rate for matching sets and non-matching sets converged at 83.4% and 83.6%.[140] Additionally, when considering reproducibility, the rates increased to 72.4% for matching sets and 54.6% for non-matching sets.[141] Thus, simply regrouping the "inconclusive" responses eliminated the Majority's concern entirely without "compromis[ing] the reliability of a conclusive determination."[142]

Moreover, Justice Gould observed that "consistent results from separate examinations of the same sample, by either the same or a different examiner, are not necessarily desirable."[143] As he found when analyzing the data, most of the instances where conclusions were changed in subsequent examinations resulted in a shift to the correct categorization.[144] For instance, a sample initially misclassified as an "identification" may have been properly reclassified as an "exclusion." While this technically created a lower repeatability or reproducibility score, the change was desirable as it resulted in a correct conclusion. Hence, Justice Gould opined that the Ames II Study's shifting of conclusions to the ground truth indicated that the

134. *Id.*
135. *Id.* at 1015 (citing PCAST REPORT, *supra* note 4, at 152).
136. *Id.* at 1015–16.
137. *See id.*
138. *See id.* at 1016–25. Justice Gould also included a more clearly stated "recap" of the Majority's entire analysis. *Id.* at 1025–27.
139. *Id.* at 1016–17, 1021–22.
140. *Id.* at 1021.
141. *Id.* at 1023, 1025.
142. *Id.* at 1021.
143. *Id.* at 1016.
144. *See id.* at 1018–25.

AFTE Theory had a high reliability. He further opined that the conclusory shifts demonstrated the level of caution that examiners exercise when conducting comparisons, confirming the functionality of the peer review process required by the firearm identification methodology.[145]

Lastly, Justice Gould addressed the Majority's critique of the lack of standards or controls governing classification of class or individualized characteristics.[146] He agreed that published standards would be "helpful," but stated that the Majority ignored "evidence that examiners make correct determinations in practice."[147] While Justice Gould acknowledged that the Majority's reasoning was not necessarily "irrational or unreasonable,"[148] he ultimately concluded that the trial court's decision to admit the testimony was reasonable and should not have been overturned under an abuse of discretion standard.[149]

Justice Gould's exhaustive dissent represents the logic underlying many proponents' arguments for admitting firearm identification testimony. Additionally, it highlights the exact reason why some courts choose to admit this evidence despite vehement protests from defense attorneys and other scholars.

D. The 2023 Amendment to FRE 702

Shortly after the *Abruquah* decision was published, in December 2023, a new amendment to FRE 702 went into effect, aimed at fixing "widespread misapplication of the Rule by courts."[150] The Advisory Committee Notes accentuate the trend of rigorously applying *Daubert* to forensic evidence and, if it is deemed admissible, the shift from definitive "match" language to limited testimony.[151] The Notes emphasize 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."[152] Specifically referencing feature-comparison evidence, the Advisory Committee comments that "opinion testimony regarding [its] weight . . . must be limited to those

145. *Id.*

146. *See id.* at 1029–31.

147. *Id.* at 1029. The evidence referred to consisted of several studies, including Ames II, which analyzed examiners' ability to formulate correct conclusions "despite the risk of subclass carryover" under the AFTE Theory. *Id.* at 1030–31.

148. *Id.* at 1031.

149. *Id.* at 1031–32.

150. Mark A. Behrens, *A Brief Guide to the 2023 Amendments to the Federal Rules of Evidence*, FEDERALIST SOC'Y: FEDSOC BLOG (Jan. 30, 2024), https://perma.cc/P7F9-D7GR.

151. *See* FED. R. EVID. 702 advisory committee's note to 2023 amendment.

152. *Id.*

inferences that can reasonably be drawn from a reliable application of the principles and methods."[153] It also stresses the importance of judges considering error rates of the utilized methodology to determine "how often the method produces accurate results."[154]

Despite the close timing of its enactment after the *Abruquah* decision, the amendment was not directly a result of the Maryland Supreme Court's analysis.[155] Rather, the amendment was under development at least as early as June 2022, when the then-chair of the Advisory Committee on Evidence Rules, Patrick Schiltz,[156] criticized judges' beliefs that "the reliability requirements set forth . . . are questions of weight and not admissibility, and more broadly that expert testimony is presumed to be admissible."[157] Although the amendment appears to place the legislature in the position of determining the admissibility of evidence, and possibly encroaching on judges' roles as gatekeepers, this is not the case. Instead, the amendment is most properly understood as enforcing judges' and attorneys' own views regarding how admissibility should be determined to ensure consistent application of reliability principles across the country.

IV. EVALUATING THE *ABRUQUAH* OUTCOME

The *Abruquah* court's ultimate decision to exclude the firearm identification evidence accurately represented a complete understanding of the necessary reliability standard for forensic science principles. However, the Majority's reasoning was not without flaws. While several of its analyses of the *Daubert-Rochkind* factors are virtually indisputable and thus do not evoke concern,[158] others warrant further discussion.

A. General Acceptance

Beginning with the factor of general acceptance, the Majority was correct to be skeptical of both parties' arguments. This factor refers to the *Frye* standard which, importantly, does not define who is included in the relevant scientific community.[159] As the *Abruquah* Majority

153. *Id.*

154. *Id.*

155. *See* Behrens, *supra* note 150.

156. Patrick Schiltz is also the Chief United States District Court Judge for the District of Minnesota. *Id.*

157. *Id.* (quoting Memorandum from Hon. Patrick J. Schiltz, Advisory Comm. on Evidence Rules, to Hon. John D. Bates, Standing Comm. on Rules of Prac. & Proc. 6 (May 15, 2022), https://perma.cc/DQX5-F4Q2).

158. These nonproblematic factors include whether the discipline is testable (factor one), whether the opinions were developed for litigation (factor six), and the expert's level of care (factor nine).

159. *See* Frye v. United States, 293 F. 1013, 1014 (D.C. Cir. 1923); *see also* PCAST REPORT, *supra* note 4, at 42 n.90.

noted, "whether the AFTE Theory . . . is generally accepted . . . is largely dependent on what the relevant community is."[160] It is true that defining the relevant community to simply consist of those who make their livelihood and benefit from its continued acceptance in court would not account for potential bias.[161] Thus, because of the ongoing dispute between proponents and others familiar with the AFTE Theory, the Majority properly weighed this factor neutrally.

B. Peer Review

Next, requiring that studies are published in peer-reviewed journals is critical to ensure that forensic science is held to a high standard—one of accountability. Although it would be ideal to only admit scientific evidence that has been deemed to be 100% accurate, that feat would be almost impossible to accomplish. At second best, then, should be the goal of ensuring that only scientific data that can be reviewed and argued against is admitted into the courtroom.

As the Majority noted, the Ames I and II Studies had yet to be peer reviewed or published at the time of the evidentiary hearings in the *Abruquah* case.[162] In January and September 2023, reports on the accuracy phase and reproducibility and repeatability phases of the Ames II Study were published.[163] To date, it does not appear that the full dataset has been released to enable public researchers to conduct verification studies.[164] As such, the Majority's weighing of this factor toward exclusion was proper and remains so, at least until the Ames II Study can be fully reviewed.

A primary reason to hold forensic scientists accountable by requiring peer reviewed publications is the prominence of the "CSI effect." Stemming largely from the mass-publicization of forensic science in the media, the CSI effect refers to the phenomenon where jurors, with no scientific or legal background, believe (1) that forensic evidence is available, or should be presented, in all cases, and (2) that admitted forensic evidence holds significant weight.[165] Allowing

160. Abruquah v. State, 296 A.3d 961, 994 (Md. 2023).

161. *Id.*

162. *Id.* at 988.

163. Thomas D. Albright & Nicholas Scurich, *A Call for Open Science in Forensics*, 121 PNAS 1, 3 (2024), https://perma.cc/73L2-JU75. For the study on accuracy, see generally Keith L. Monson et al., *Accuracy of Comparison Decisions by Forensic Firearms Examiners*, 68 J. FORENSIC SCIS. 86 (2023) [hereinafter Monson, *Accuracy of Comparison Decisions*], https://perma.cc/46L8-FARX. And for the study on reproducibility and repeatability, see generally Keith L. Monson et al., *Repeatability and Reproducibility of Comparison Decisions by Firearms Examiners*, 68 J. FORENSIC SCIS. 1721 (2023), https://perma.cc/5SGV-GNJW.

164. *See* Albright & Scurich, *supra* note 163, at 3; *see also* Kori Khan & Alicia L. Carriquiry, *Shining a Light on Forensic Black-Box Studies*, 10 STAT. & PUB. POL'Y 1, 7 (2023), https://perma.cc/QPU3-9FUJ.

165. *See* Donald E. Shelton, *The 'CSI Effect': Does It Really Exist?*, NAT'L INST. JUST. (Mar. 16, 2008), https://perma.cc/B9U6-VBMK. *See generally* John

testimony that a particular "unknown" bullet came from the defendant's firearm is damning, and it subverts the jurors' unbiased fact-finding duty. While merely limiting testimony to "consistent with" may not seem that different, it at least (in theory) allows the jury to draw their own conclusions from the evidence, rather than be told with absolute certainty what occurred.

C. Error Rate and Relationship Between Reliability of Methodology and Opinion

That said, the Majority confuses the matter by weighing the "known or potential error rate" and "relationship between reliability of methodology and opinion" factors in favor of exclusion.[166] The Majority's issues with the Ames II Study regarding accuracy, repeatability, and reproducibility all derive from the "inconclusive" categories. To that end, Justice Gould made compelling arguments about focusing on conclusive determinations and pooling together the "inconclusive" categories, both of which demonstrated the reliability of firearm identification.

While it is technically correct that "inconclusive" responses are still somewhat harmful to defendants when used as evidence in court, it is likely less common for the government to present evidence that does not tend to affirmatively inculpate the defendant. Really, an "inconclusive" result being presented hurts the party presenting it just as much as it helps them. For example, as a defendant, while an "inconclusive" result shows the government cannot say it *was* you, it also does not say it *wasn't*. Thus, it seems proper to focus on the Ames II results that deal with only conclusive responses (i.e., "identification" or "exclusion"). As Justice Gould noted, this approach aligns with the suggestions of PCAST.[167] Further, even under this approach, the error rates remained well under the PCAST-recommended threshold of 5.0%, while the reproducibility and repeatability rates greatly improved.[168]

Importantly, the Majority completely ignored the fact that, in most instances when a subsequent examination reached a different conclusion, the second response aligned with the correct result. This demonstrates that firearm examiners often make correct conclusions initially and do not frequently disagree with prior correct responses,

---

Alldredge, *The "CSI Effect" and Its Potential Impact on Juror Decisions*, 3 THEMIS: RSCH. J. JUST. STUD. & FORENSIC SCI. 114 (2015).

166. *See Abruquah*, 296 A.3d at 991, 997.

167. *See id.* at 1014 (Gould, J., dissenting).

168. *See id.* at 1016, 1026; *see also supra* Section III.C. *But see* Nicholas Scurich et al., *The Hawthorne Effect in Studies of Firearm and Toolmark Examiners*, 70 J. FORENSIC SCIS. 1329, 1334–35 (2025), https://perma.cc/XJD2-9PKU (finding that firearm examiners made more decisions of "inconclusive" when they knew they were being tested than in normal casework, calling into question the accuracy of past studies' error rates).

supporting the reported low error rates. Moreover, it also demonstrates that examiners tend to catch erroneous conclusions, meaning the verification methodology of comparisons is effective in this context. So, when referring to the reliability of the results in front of the *Abruquah* court, Ames I and II arguably demonstrated that the AFTE Theory's methodology supported "identification" conclusions.

D. Standards and Controls, Unjustified Extrapolation, and Alternative Explanations

Nevertheless, the Majority accurately reasoned that the lack of existing standards and controls strongly supports exclusion. This is perhaps the most important *Daubert* factor. Thus, even though the reliability of the Ames I and II Studies—and the discipline itself—is greater than the Majority credits, the weight of this factor still outweighs the factors favoring admission.[169] Considering the serious consequences of identification testimony, which could result in the significant taking of a person's liberty or even life, it is crucial that examiners only testify to what science can support. To be able to testify to such conclusive evidence as, essentially, "the defendant committed the crime," the process and methodology followed to reach that conclusion should be well-established. But the AFTE Theory can in no way be called well-established.

Although most forensic feature-comparison disciplines rely on subjective interpretations to some extent, the AFTE Theory has almost no empirical method. The Theory itself explicitly states that "the interpretation of individualization/identification is subjective in nature" and relies on examiners' own understanding of "sufficient agreement" based on their experience and training.[170] This is problematic in several regards. First, most examiners do not begin work having relevant experience, and their preparatory training may be minimal. Without detailed processes from which to draw conclusions, firearm examiners are left to their own devices in conducting evidentiary comparisons.

---

169. The Majority stated that it did "not question that firearms identification is generally reliable . . . in identifying whether patterns or markings on 'unknown' bullets or cartridges are consistent or inconsistent with those on bullets or cartridges known to have been fired from a particular firearm." *Abruquah*, 296 A.3d at 996. This indicates that the current issue in the field seems to be more of establishing better standards and methodologies that support making "identification" conclusions—such as addressing the issue of subclass versus individualizing characteristics—rather than necessarily increasing reliability. Perhaps, though, once this is done and more is known about the relationship between subclass and individualizing characteristics, it may be found that the reliability is not as high as indicated by the Ames or previous studies.

170. *Theory of Identification as It Relates to Toolmarks*, *supra* note 41.

Second, these types of comparisons are difficult to perform even for examiners who have ample experience. Just considering Justice Gould's adjusted repeatability and reproducibility rates from the Ames II Study, firearm examiners disagreed with themselves around 20% of the time and with each other about 30–45% of the time.[171] Without a proper methodology to follow regarding how to classify class, subclass, and individualizing characteristics, how many features to compare, or how many similarities or differences are required to form certain conclusions, examiners are not going to consistently perform the same analyses or reach the same results.

This lack of structured procedure increases the likelihood that examiners will reach different conclusions for the same evidence. Moreover, it raises the concern that "identification" conclusions will be formed from too few comparison points or from impressions of insufficient quality to assure reliability. Keep in mind that, currently, the AFTE Theory does not even list a minimum number of reference points that must be compared and found to be similar to support an "identification" conclusion. Because firearms are mass-produced by highly specialized and precise machinery, there are likely to be extensive similarities in the manufacturing-related markings, leaving little to distinguish between weapons. Therefore, it seems implausible that a firearm examiner can even identify these minute differences, let alone distinguish between them. As a result, most of the details that examiners do focus on and compare are more likely to be subclass characteristics than individual characteristics. And these themselves cannot support an "identification" conclusion.

Consequently, and consistent with the Majority's holding, there is an analytical gap between the AFTE Theory and the ability to identify a specific firearm as the source of a fired bullet or cartridge case. As it stands, firearm examiners do not know the extent of the similarities, or importantly, differences, between firearms that are produced in the same batch at the same factory.[172] Realistically, every firearm produced by a specific factory location will share the majority of their characteristics.[173] Yet, the identification theory largely hinges on the existence of minute differences in every firearm produced by the ever-so-slight, gradual wear-and-tear of the manufacturing machines with the production of each unit. However, there has not been research done regarding the similarities of sequentially produced firearms to even know that there is a cognizable difference in the markings they would produce on fired bullets or cartridge cases. As such, firearm examiners who testify to an "identification" are necessarily not accounting for obvious explanations since they do not have the ability to test all other firearms that were produced in

171. *See Abruquah*, 296 A.3d at 1017, 1019, 1021, 1023, 1025 (Gould, J., dissenting).

172. *See supra* notes 36–38, 76–77 and accompanying text.

173. *See supra* notes 36–38 and accompanying text.

the same batch to determine what details were individualizing versus simply subclass characteristics.

Moving forward, it is true that the development and advancement of artificial intelligence and computer programming may enable less subjective analyses of bullets and cartridge cases.[174] In turn, this could allow for more accurate classification of impression details and result in more reliable comparisons and conclusions. However, that time has yet to come.[175]

Ultimately, while *Daubert* and FRE 702 require trial judges to act as gatekeepers and only admit reliable expert testimony, the *Abruquah* dissenters had a compelling argument that the degree of reliability should be a question for the jury—after all, the jury is tasked with determining what to believe and applying that to the case to deliver a verdict.[176] Given Justice Gould's breakdown of the Ames II data into promising statistics of repeatability and reproducibility, in addition to the low false-positive rates indicating accuracy, the reliability of firearm identification evidence (at its current state)

---

174. In fact, the *Abruquah* court briefly mentioned the more objective "consecutive matching striae" (CMS) method but dismissed it because that technique had not been utilized in the case before the court. *Abruquah*, 296 A.3d at 975 n.9. The CMS method is typically used in conjunction with visual assessment—examiners will utilize a procedure akin to the AFTE Theory to derive individualizing characteristics before employing CMS to "quantify the similarity of two lands." Susan Vanderplas et al., *Comparison of Three Similarity Scores for Bullet LEA Matching*, 308 FORENSIC SCI. INT'L 1, 1 (2020). The National Institute of Justice has called CMS "simply a means of articulating the best known nonmatch described and defined by the AFTE Theory of Identification." *Firearms Examiner Training: Pattern Identification*, NAT'L INST. JUST. (2025), https://perma.cc/F9FM-9KBX.

In other words, the CMS methodology is employed after using visual comparisons merely to determine the level of agreement between an "unknown" and "known" sample. However, it is important to recognize that (1) the CMS methodology still relies extensively upon examiners' subjective analyses and classifications of characteristics as class versus subclass versus individualizing to produce accurate results; and (2) still relies on examiners' subjective interpretations of whether the striations examined are in agreement. *See id.*

175. *See* Alan H. Dorfman & Richard Valliant, *Inconclusives, Errors, and Error Rates in Forensic Firearms Analysis: Three Statistical Perspectives*, 5 FORENSIC SCI. INT'L: SYNERGY 1, 4 (2022) (discussing the future of using machine learning for bullet and cartridge case comparisons).

176. There was a second dissent in the *Abruquah* case, by Justices Hotten and Eaves. *See Abruquah*, 296 A.3d at 998–1005 (Hotten, J., dissenting). That dissent focused primarily on the Majority misinterpreting *Daubert* and improperly side-stepping the purpose of the jury. *See id.* at 1000–02. The Justices opined that the circuit court correctly categorized the defendant's objection as "more suited to the *weight* such evidence should be given *at trial*" rather than the outright admissibility of the testimony. *Id.* at 1002. However, as mentioned in Section III.D *supra*, the 2023 FRE 702 amendment was partly driven as a direct response to and rejection of these beliefs by judges.

should not be the limiting factor. Instead, the hurdle that must be overcome is the standardization of analysis and the production of detailed methodology for firearm examiners to follow that better accounts for mass-produced similarities. While the procedure need not be completely objective, citizens, attorneys, and judges cannot rely on thousands of forensic examiners across the country to make consistent findings when they have no concrete guidelines to follow.

V. MOVING FORWARD (OR BACKWARD?): THE FUTURE OF FORENSIC FEATURE-COMPARISON EVIDENCE IN COURTROOMS

Despite reaching the proper conclusion, the *Abruquah* decision exemplifies the minority approach to the admissibility of firearm identification evidence. This Part explores the scope of the holding in terms of what evidence, if any, it excludes from Maryland courts. Then, it discusses how other states' courts have used *Abruquah* both to admit and exclude similar evidence. Finally, it addresses the implications that the reasoning employed by the *Abruquah* Majority may have on the admissibility of other feature-comparison evidence.

A. The Scope of the Abruquah *Decision Within Maryland*

From afar, the *Abruquah* decision appears to instill a harsh prohibition on firearm examiners identifying a firearm as the source of any ballistics evidence.[177] However, two important caveats on the actual breadth of the holding emerge upon closer review.

First, given the limiting language about its consideration of only those studies presented to the circuit court, the *Abruquah* court clearly did not intend to bestow "a final verdict on the reliability of firearm identification evidence."[178] In fact, since the *Abruquah*

177. Interestingly, a recent case from the Appellate Court of Maryland exhibits the impact of *Abruquah* beyond "identification" conclusions. *See* Bryant v. State, No. 1786, Sept. Term, 2023, 2025 WL 2398791 (Md. App. Ct. Aug. 19, 2025). In that case, a firearm examiner testified that bullets and cartridge cases from two shooting scenes "were consistent with each other" and the same "unknown" firearm "could not be excluded as the source" of all evidence. *Id.* at *10. Relying on the abuse of discretion standard, the court ruled that the testimony was not erroneously admitted. *Id.* at *9–10. The court emphasized that "firearms identification is generally reliable . . . in identifying whether patterns . . . are consistent or inconsistent." *Id.* at *9 (quoting *Abruquah*, 296 A.3d at 996). It held that there was no "analytical gap" between the examiner's methodology and conclusion because the comparison of class characteristics alone could "reliably support" making exclusions. *Id.* at *10 (quoting *Abruquah*, 296 A.3d at 970). Thus, despite the examiner's reliance on the AFTE Theory's methodology, the fact that the conclusion was not an "identification" that would require the exclusion of "all alternative sources" was highly relevant to the court. *Id.*

178. Stevens v. State, No. 85, Sept. Term, 2023, 2024 WL 5066043, at *10 (Md. App. Ct. Dec. 11, 2024), *cert. denied*, 334 A.3d 830 (Md. 2025).

decision, it appears that lower Maryland courts have gone on to consider several additional studies and have still upheld the prohibition.[179] For example, the Appellate Court of Maryland contemplated nineteen new studies cited by the State in an attempt to show that the *Abruquah* holding had been materially displaced.[180] However, the court held that the studies did nothing to demonstrate the reliability of the AFTE Theory since none of them tested reproducibility or repeatability.[181] Moreover, their double-digit false-positive rates continued to showcase the Theory's low accuracy.[182]

More recently, on July 21, 2025, the Appellate Court issued a decision in *Dunbar v. State*,[183] reversing the convictions of two defendants "based on the erroneous admission of expert firearms identification testimony."[184] In that case, the *Abruquah* decision had come out after the defendants' trials but before their direct appeal was decided.[185] On appeal, the State relied on the *Abruquah* court's couching of its holding as limited to the studies presented at the *Daubert-Rochkind* hearing.[186] Given that the defense had not requested such a hearing, the State argued that it was not afforded the ability to offer additional studies so as to support the testimony it presented.[187] Ultimately, the court rejected this argument, finding plain error in "[t]he admission of . . . conclusive testimony matching the recovered projectiles to certain weapons."[188] Indeed, even if the court had considered additional studies, it is unlikely that the result would have changed for the same reasons as stated in *Harris*.

As a second large caveat, the *Abruquah* court did not clearly specify whether its holding encompassed bullets and cartridge cases or just bullets. Although the conclusion itself explicitly referred to "bullets,"[189] with no mention of cartridge cases, the Majority could have been referring to both cartridge case and bullet comparisons throughout most of the opinion. For instance, the methodology critiqued by the Majority, the AFTE Theory of Identification, does not explicitly mention bullets or cartridge cases and is utilized when examining and comparing either.[190] Additionally, the studies the

179. *See, e.g.*, Harris v. State, No. 2209, Sept. Term, 2022, 2024 WL 4161704, at *7–8 (Md. App. Ct. Sept. 12, 2024), *cert. denied*, 330 A.3d 660 (Md. 2025).

180. *Id.* at *7.

181. *Id.* at *8.

182. *Id.*

183. Nos. 611, 639, 1598, Sept. Term, 2023, 2025 WL 2027549 (Md. App. Ct. July 21, 2025).

184. *Id.* at *1.

185. *Id.* at *16.

186. *Id.* at *19.

187. *Id.*

188. *Id.* at *21.

189. Abruquah v. State, 296 A.3d 961, 998 (Md. 2023).

190. *See Module 09: Cartridge and Shotshell Examination, AFTE Theory of Identification*, NAT'L INST. JUST. (2025), https://perma.cc/F39G-VP3K; *Module 11:*

Majority referenced entailed bullet and cartridge case comparisons, and the PCAST Report merely referred to Ames I generally.[191] This begs the question whether Maryland courts going forward will allow admission of firearm identification testimony relating to cartridge cases or whether the state-wide ban on identifying a firearm as a source is applicable to both types of evidence.

In reviewing the case law from lower Maryland courts since *Abruquah*, there seems to be an understanding that the decision prohibits admission of firearm identification testimony relating to both bullets and cartridge cases.[192] This is logical given that the Ames II Study demonstrated that they had similar false-positive and false-negative rates—the false-positive rates were estimated at 0.656% and 0.933% while the false-negative rates were estimated to be 2.87% and 1.87% for bullets and cartridge cases, respectively.[193] But, there could be a meaningful distinction between the admissibility of opinions based on bullets versus cartridge cases that is not currently known.[194]

B. The Broader Implications of the Abruquah *Decision*

*1. Post-*Abruquah *Rulings on Firearm Identification Evidence in Other States*

Since *Abruquah* was decided, courts in other states have cited the case, some to hold similarly and others to distinguish. Like the Appellate Court of Maryland, other courts have considered new or additional studies upon challenges by a defendant to firearm

---

Bullet Comparison and Identification, AFTE Theory of Identification, NAT'L INST. JUST. (2025), https://perma.cc/P5Q4-MNQT.

191. *See Abruquah*, 296 A.3d 961; PCAST REPORT, *supra* note 4, at 110–14.

192. *See, e.g.*, Stevens v. State, No. 85, Sept. Term, 2023, 2024 WL 5066043, at *8 (Md. App. Ct. Dec. 11, 2024) (writing that the State itself argued that the firearm examiner could testify "'with the sole exception of offering his ultimate conclusion that' the cartridges entered into evidence were fired with the firearm that he examined"), *cert. denied*, 334 A.3d 830 (Md. 2025); Harris v. State, No. 2209, Sept. Term, 2022, 2024 WL 4161704, at *8 (Md. App. Ct. Sept. 12, 2024) (holding the trial judge abused its discretion in allowing a firearm examiner to testify that cartridge cases were fired by a specific firearm), *cert. denied*, 330 A.3d 660 (Md. 2025).

193. *See* Monson, *Accuracy of Comparison Decisions*, *supra* note 163, at 97.

194. One difference could arise from the quantity and quality of the impressions being compared on each type of evidence. As a nature of the mechanical action of the firearm during discharge, more points of reference are imprinted on the cartridge case than the bullet. For example, breech face, firing pin impressions, and random impressions are imprinted on the cartridge case, whereas the bullet only acquires striations from moving down the barrel. *See* THE ROYAL SOC'Y, *supra* note 18, at 31; *see also supra* Section I.A. From this, it may be determined that comparisons of cartridge cases are more individualizing than those of bullets, such that they become admissible while bullets remain excluded.

identification evidence. For example, in *State v. Raynor*,[195] a Connecticut Superior Court considered twenty-one peer-reviewed studies that had been published after the PCAST Report.[196] It found that the studies produced error rates much below the suggested 5.0% threshold and that there was no issue with excluding "inconclusive" responses as errors.[197] Based on this, the court reached the opposite holding as Maryland courts on several of the *Daubert* factors, ruling that examiners could testify to a "reasonable degree of certainty," which is a stronger conclusion than allowed under *Abruquah*.[198]

Even without relying on new studies, some courts have explicitly distinguished the *Abruquah* holding. In *People v. Rodriguez-Ortiz*,[199] the Colorado Court of Appeals held that the "trial court did not abuse its discretion by finding that the prosecution's expert's firearms toolmark analysis satisfied the [statute's] threshold of baseline reliability and that any shortcomings went to the weight of the evidence and not its admissibility."[200] The court differentiated *Abruquah* in two ways. First, it reasoned that Colorado's reliability standard did not demand analysis of all ten factors as required by the *Daubert-Rochkind* standard in Maryland.[201] Second, because the examiner testified that "his determinations were not based on any quantitative analysis, and that his final conclusion was a subjective determination, not an objective one," the opinion was "qualified" and distinct from the "*unqualified* opinion" in *Abruquah*.[202]

In contrast, on May 29, 2025, the Oregon Court of Appeals held that the State had failed to meet its burden to show that the AFTE Theory was scientifically valid or reliable and ruled that the admission of the firearm identification evidence was not harmless.[203] This case involved the same defendant as *United States v. Adams*,[204] where the federal District Court of Oregon drastically limited the testimony of the firearm examiner to only class characteristics after analyzing the *Daubert* factors and concluding that the AFTE methodology "lack[ed] . . . any objective standard" to make it a

195. No. HHD-CR13-0667367, 2024 WL 3579515 (Conn. Super. Ct. Apr. 16, 2024).

196. *Id.* at *7.

197. *Id.* at *7–10.

198. *Id.* at *13.

199. 574 P.3d 1196 (Colo. App. 2025).

200. *Id.* at 1210.

201. *Id.* at 1209.

202. *Id.* at 1210. Just over a month later, the Colorado Court of Appeals again ruled in contrast to *Abruquah*. *See* People v. Alvarado-Vasquez, No. 23CA1491, 2025 WL 2170561, at *9–11 (Colo. App. July 31, 2025) (finding no abuse of discretion in the admission of firearm identification testimony, based partly on Justice Hotten's dissent in *Abruquah*).

203. State v. Adams, 572 P.3d 291, 296, 314 (Or. Ct. App. 2025).

204. 444 F. Supp. 3d 1248 (D. Or. 2020).

"*scientific* inquiry."[205] The Oregon Court of Appeals ultimately agreed, holding that the AFTE Theory was not scientific and thus could not be helpful to the jury when presented as science.[206]

In coming to this decision, the Oregon Court of Appeals relied heavily on the federal judge's findings, accepting his analysis of the peer review, standards and control, and general acceptance factors.[207] This enabled the court to primarily focus on the circularity and non-instructiveness of the AFTE Theory: "[T]he method does not actually *measure* the degree of correspondence between shell cases or bullets; rather, the practitioner's decision on whether the degree of correspondence indicates a match ultimately depends entirely on subjective, unarticulated standards and criteria arrived at through the training and individualized experience of the practitioner."[208] It disagreed with the federal court, and the *Abruquah* Majority, on the issue of error rate, holding that the factor was merely neutral.[209] As the Oregon court only referenced *Abruquah* once,[210] its decision demonstrates that the exclusion of identification testimony may be an upcoming trend across the nation as individual judges begin to apply *Daubert* and other reliability factors to firearm comparison evidence.

2. The Abruquah *Reasoning's Possible Effect on the Admissibility of Other Forensic Feature-Comparison Evidence*

Examining the *Abruquah* holding is critical to understanding the direction courts may be heading in their admission of forensic evidence generally. With the 2023 FRE amendment and courts applying the *Daubert* standard more rigidly, it is possible that courts will not only begin to exclude firearm identification evidence, but also other types of forensic feature-comparison evidence. This Section briefly discusses the possible implications of *Abruquah* on two other forensic disciplines, each of which relates to firearm evidence in a distinct way: latent fingerprints and footwear and tire impressions.

a. Latent Fingerprints

Latent fingerprint and firearm identification evidence are similar in many ways. For instance, they both have a long-standing and frequent utilization in court to help identify a defendant as the

205. *Id.* at 1261, 1266–67. In the federal case, Judge Mosman held that the AFTE Theory was not reproducible, had a misleadingly low error rate, was not sufficiently peer reviewed, and was not generally accepted within the scientific community. *Id.* at 1264–67. The only *Daubert* factor Judge Mosman found to be met was the existence of standards and controls. *Id.* at 1266.

206. State v. Adams, 572 P.3d at 311–12.

207. *Id.* at 307–08.

208. *Id.* at 295.

209. *Id.* at 311–12.

210. *See id.* at 298 n.8.

perpetrator of a crime.[211] Moreover, both disciplines have faced extensive criticisms, largely stemming from their subjective methodology and unknown reliability.[212] However, unlike the lackadaisical effort from firearm examiners following the NAS Report, fingerprint analysts acted quickly to conduct studies to support their work's reliability.[213] In sharp contrast to its findings on firearm evidence, the subsequent PCAST Report noted that two of the new properly-designed latent print studies demonstrated "a foundationally valid subjective methodology."[214]

As with firearm evidence, even after the publication of the PCAST Report it was only recently that a few judges began to question the admissibility of latent fingerprint evidence.[215] But even then most courts still tended to admit the evidence without any rigorous reliability examination.[216] While fingerprint studies arguably better support the discipline's reliability than do the firearm identification studies, it is possible that fingerprint evidence's admissibility is at risk collaterally. Because of the prevalence at which fingerprints are used to implicate and convict defendants, and because of their heightened reliability in the eyes of the jury,[217] the use of fingerprint identifications very well could have the same large-scale negative effect as the admission of firearm evidence—the deprivation of a defendant's right to a fair trial. For this reason, as courts more commonly restrict firearm evidence, they could become

211. NAS REPORT, *supra* note 10, at 43.

212. PCAST REPORT, *supra* note 4, at 89. Fingerprint comparison uses the "ACE-V" methodology, which requires examiners to subjectively *analyze* the quality of the impression to determine its suitability for comparison, *compare* the "unknown" and "known" impressions to determine the level of agreement between them, *evaluate* that level of agreement to form a conclusion, and have another independent examiner *verify* the conclusion. *Id.* at 89 n.248.

213. *Id.* at 95.

214. *Id.* at 101. Despite this, the PCAST Report also stated that the discipline was not valid as applied, and because of the studies' very inconsistent false-positive rates, it was unclear whether either accurately represented the error rate in casework. *Id.* at 97, 101.

215. *See* Brandon L. Garrett, *The Reliable Application of Fingerprint Evidence*, 66 UCLA L. REV. DISCOURSE 64, 66 (2018); Justin K. Fahringer, *Advancing Forensic Science in Criminal Justice: Latent Print Evidence—A Case Study*, CHAMPION, June 2022, at 34, 34.

216. *See* Garrett, *supra* note 215, at 66–68 (discussing how judges have even held that fingerprint evidence was reliable by virtue of "the use of the technique for 'almost a century'" (quoting United States v. Baines, 573 F.3d 979, 990 (10th Cir. 2009))).

217. *See* Kimberly Schweitzer & Narina Nuñez, *What Evidence Matters to Jurors? The Prevalence and Importance of Different Homicide Trial Evidence to Mock Jurors*, 25 PSYCHIATRY PSYCH. & L. 437, 443, 447 (2018) (examining several studies that surveyed what evidence jurors believed was most important and finding that fingerprint evidence was generally rated second or third, only behind DNA and video confession evidence).

stricter in their admission of fingerprint evidence and begin to exclude it to protect defendants. More likely, courts will simply limit the testimony to a degree of similarity rather than identification.[218]

b. Footwear and Tire Impressions

Footwear and tire impression evidence are similar to firearm identification evidence in a different way—all three are created by items that are mass produced. As a result, analysis of each relies on the classification and comparison of class, subclass, and individualizing characteristics.[219] Because the *Abruquah* Majority's discussion focused largely on the unknown reliability of identifying and comparing individualizing characteristics, courts moving forward may be inclined to exclude footwear and tire comparison evidence on similar grounds.

This is especially true considering that an important distinction between footwear and tire impressions and firearm impressions is the consistency in their creation. Firearm impressions are formed in a highly consistent manner by virtue of always being deposited on the rigid but malleable metals that make up bullets and cartridge cases. On the other hand, footwear and tire impressions are frequently made in less firm substrates, such as soil, mud, and snow. Because of the pliability of these materials, it is common for details of a shoe or tire impression to be less clearly defined or even become distorted from the source.[220] As a result, tires and shoes produce less predictable and less reproducible impressions than those made by firearm components.

However, because of this known variability and the resultant reduced reliability in comparisons, courts are currently more likely to exclude footwear and tire evidence. This is also partly because fewer studies have been conducted regarding footwear and tire impression

218. *See* U.S. DEP'T OF JUST., UNITED STATES DEPARTMENT OF JUSTICE UNIFORM LANGUAGE FOR TESTIMONY AND REPORTS FOR THE FORENSIC LATENT PRINT DISCIPLINE 2 (2020), https://www.justice.gov/media/1072031/dl (allowing "source identification" as an acceptable conclusion for latent fingerprint examiners to testify to).

219. Similarly to firearm comparison, footwear examinations begin by comparing class characteristics, which would be the size of the shoe/impression, the design or tread, and the wear. PCAST REPORT, *supra* note 4, at 114; *see* U.S. DEP'T OF JUST., UNITED STATES DEPARTMENT OF JUSTICE UNIFORM LANGUAGE FOR TESTIMONY AND REPORTS FOR THE FORENSIC FOOTWEAR DISCIPLINE 2 (2020), https://www.justice.gov/olp/page/file/1284771/dl (explaining "outsole design," "physical size," and "wear"). Then, if the class characteristics are consistent, the examiner considers "randomly acquired characteristics (RACs)," which are "marks on a shoe caused by cuts, nicks, and gouges in the course of use." PCAST REPORT, *supra* note 4, at 114.

220. *See* Yaron Shor et al., *Inherent Variation in Multiple Shoe-Sole Test Impressions*, 285 FORENSIC SCI. INT'L 189, 189 (2018).

evidence,[221] which makes it more difficult for a court to rule properly on their reliability and admissibility. Thus, since this type of evidence is already more frequently excluded and is less likely to weigh as heavily with jurors, judges may not become overly cautious about it and there may only be a minimal reduction in its admissibility.

CONCLUSION

A driving factor in the trend of prohibiting firearm identification testimony is likely the prevalence of gun crime and the sheer volume of firearm-based offenses. Allowing mildly reliable evidence may not seem too worrisome when it involves less common forensic evidence, such as footwear and tire impressions. However, when so many prosecutions are of firearm crimes, and many cases *only* have firearm evidence, the cost of allowing identification testimony is too great. Because of the increased value on examiners' conclusions in those cases, it is crucial that courts require them to be based on reliable principles. Otherwise, consistently admitting this high value and, essentially, damning evidence offers alleged criminals no real chance to defend themselves and effectively violates their right to a fair trial.

Consequently, the *Abruquah* Majority made the correct decision in holding that firearm identification testimony does not meet the standard set forth by *Daubert*. Courts should only reconsider admitting identification conclusions in the future if a more detailed methodology is created that accounts for the differences between simple comparisons and comparisons of mass-produced items. Until then, juries are best left being presented with more probabilistic conclusions such as "consistent with." Still, if that is observed to carry the same weight as definitive conclusions, then testimony should be further limited to only pertain to class characteristics.

Melissa Stuckey[*]

---

221. In the PCAST Report, footwear "identification" conclusions were noted to be "breathtaking—but lack scientific foundation," because no black-box studies had determined the validity of the discipline, and no other studies demonstrated its reliability or accuracy. PCAST REPORT, *supra* note 4, at 115–17.

* J.D., 2026, Wake Forest University School of Law; B.S. Forensic and Investigative Science, 2023, West Virginia University. I am grateful to Dr. Alyse Bertenthal and Professor Ron Wright for their ongoing guidance and support, both with this Note and throughout my law school experience, as well as Dr. Casper Venter for his mentorship in my forensic science education. Many thanks also to Cullen Irwin, Palmer Dayhuff, Nick Walters, Grace Mohlin, Vivian Bolen, Will Boyce, and the rest of the *Wake Forest Law Review* for their assistance in making this piece publishable. Finally, thank you to Jocelyn Mayo for enhancing my understanding of one of the only forensic science classes I did not take at WVU—Firearms Examination.