IN THE THIRTY-FIRST JUDICIAL CIRCUIT
GREENE COUNTY, MISSOURI

STATE OF MISSOURI,

Plaintiff;

vs.

SCOTT A. GOODWIN-BEY,

Defendant.

Case. No. 1831-CR04590
Division 6

ORDER DENYING DEFENDANT'S REQUEST TO EXCLUDE FIREARM AND TOOLMARK EXPERT TESTIMONY

Defendant Scott A. Goodwin-Bey opposes the introduction of any firearm toolmark identification evidence at his upcoming jury trial. His first motion on that topic, *Defendant's Motion to Exclude the State's Request for Admission of Ballistics and Toolmark Identification*, was filed on April 14, 2023. In that motion, he argued that rulings in his favor in a prior case (which prompted a dismissal by the State shortly before trial)are binding on this Court under the collateral-estoppel doctrine or the law-of-the-case doctrine. Those arguments were addressed in the *Order on Defendant's Motions Argued May 18, 2023*, filed on May 19, 2023.

Defendant's second motion on the topic is *Defendant's Motion for a Frye/Daubert Hearing to Determine the Admissibility of Firearms Identification Evidence to be Offered by the State*, which was filed on October 24, 2023. While no longer arguing that rulings from the prior case are binding here, Defendant urges this Court to independently reach the same result. In short, his second motion argues that "firearms toolmark identifications are not a science, but at best an art form," and that toolmark

identifications "lack any scientific basis, and should be excluded from testimony in the upcoming trial." (Defendant's Motion ¶23).

A full-day hearing on this issue was held at Defendant's request on August 8, 2024. At that hearing, Defendant appeared in person and by his attorneys, Christopher Hatley and Hannah Kahn, and the State appeared by attorneys Joshua Harrel and Kimberley Pulley. Defendant called Dr. Nicholas Scurich, Ph.D., to testify. The State then called Dr. James Hamby, Ph.D., as well as the individuals from the Missouri State Highway Patrol Crime Laboratory who performed the comparison of expended ammunition components at issue. Along with their testimony, this Court was presented with evidence including several studies, academic papers, and other exhibits. Attorneys for both sides presented oral argument and provided the Court with caselaw to review. Having now fully reviewed the evidence, arguments, and authorities presented, this Court now respectfully denies Defendant's request to exclude the toolmark-identification evidence and will allow expert testimony on the topic consistent with Missouri law and subject to the limitations of the DOJ testimony standards.

Legal Standard

In Missouri, the admissibility of expert testimony is governed by statute under Section 490.065 of the Revised Statutes of Missouri. "Prior to 2017, Section 490.065 applied a standard for the admissibility of expert testimony similar to that found in *Frye v. United States*, 293 F. 1013, 1014 (D.C. Cir. 1923)." *State v. Suttles*, 581 S.W.3d 137, 146 (Mo. App. E.D. 2019). But as of August 28, 2017, the language of Section 490.065 now mirrors Federal Rules of Evidence 702–703. In its present form, Missouri's statute now provides the following standard for expert testimony in jury trials:

(1) A witness who is qualified as an expert by knowledge, skill, experience, training, or education may testify in the form of an opinion or otherwise if:

(a) The expert's scientific, technical, or other specialized knowledge will help the trier of fact to understand the evidence or to determine a fact in issue;

(b) The testimony is based on sufficient facts or data;

(c) The testimony is the product of reliable principles and methods; and

(d) The expert has reliably applied the principles and methods to the facts of the case;

(2) An expert may base an opinion on facts or data in the case that the expert has been made aware of or personally observed. If experts in the particular field would reasonably rely on those kinds of facts or data in forming an opinion on the subject, they need not be admissible for the opinion to be admitted. But if the facts or data would otherwise be inadmissible, the proponent of the opinion may disclose them to the jury only if their probative value in helping the jury evaluate the opinion substantially outweighs their prejudicial effect;

(3) (a) An opinion is not objectionable just because it embraces an ultimate issue.

(b) In a criminal case, an expert witness shall not state an opinion about whether the defendant did or did not have a mental state or condition that constitutes an element of the crime charged or of a defense. Those matters are for the trier of fact alone;

(4) Unless the court orders otherwise, an expert may state an opinion and give the reasons for it without first testifying to the underlying facts or data. But the expert may be required to disclose those facts or data on cross-examination.

RSMo § 490.065.2. While this standard is statutory, it is still useful to consider the factors addressed in *Daubert v. Merrell Dow Pharm., Inc.*, 509 U.S. 579, 593–94 (1993), because the mirror-image language in the federal rules has been "interpreted under *Daubert* and its progeny." *Suttles*, 581 S.W.3d at 146. The Daubert factors include:

> (1) whether the expert's technique or theory can be or has been tested; (2) whether the technique or theory has been subject to peer review and publication; (3) the known or potential rate of error of the technique or theory when applied and the existence and maintenance of standards and controls; and (4) whether the technique or theory has been generally accepted in the scientific community.

State ex rel. Gardner v. Wright, 562 S.W.3d 311, 317 (Mo. App. E.D. 2018). A host of federal cases following Daubert have considered these factors in various contexts. "Importantly, however, while these cases provide relevant and useful guidance, the *Daubert* factors themselves are not controlling[.]" *Suttles*, 581 S.W.3d at 147 (citing *Wright*, 562 S.W.3d at 318).

Even among cases where the federal rules and Daubert have more direct application, courts have emphasized that "Rule 702 clearly is one of admissibility rather than exclusion," such that "expert testimony should be admitted if it advances the trier of fact's understanding to any degree." *Johnson v. Mead Johnson & Co.,* 754 F.3d 557, 562 (8th Cir. 2014) (internal citations omitted). Some federal courts have accordingly employed a simplified three-part test in applying Rule 702. *See Johnson,* 754 F.3d at 561 ("The screening requirement of Rule 702 has been boiled down to a three-part test…."). Missouri courts have followed suit after the 2017 amendment to our statute. "Since the amendment of Section 490.065, Missouri courts have used the condensed three-part standard examined in *Johnson* and discussed in *Wright* to determine the admissibility of expert

opinion testimony." *Suttles*, 581 S.W.3d at 147. That three-part standard considers (1) whether the expert is **qualified**, (2) whether the testimony is **relevant**, and (3) whether the testimony is **reliable**. *Id.*

Analysis

Defendant's primary challenge to toolmark-identification evidence falls under the third prong of that condensed standard, contending that any such evidence is not reliable enough. The *Daubert* reliability factor Defendant most focus on is the third, questioning the known or potential rate of error for toolmark examination. Put in terms of the actual statute, Defendant's argument is essentially that testimony about toolmark analysis is not "the product of reliable principles and methods." RSMo § 490.065.2(1)(c). But Defendant's position—even if otherwise persuasive—is directly at odds with recent Missouri precedent specifically addressing the admissibility of firearm and toolmark identification under the current statute. *See State v. Boss*, 577 S.W.3d 509 (Mo. App. W.D. 2019).

In *Boss*, the Western District of the Court of Appeals considered extensive testimony on the theory of toolmark identification established by the Association of Firearm and Tool Mark Examiners ("AFTE"). As was also the case here, the State's expert explained at length the process for examining the class, subclass, and individual characteristics of cartridge casings and bullets, and the criteria for making a classification of "identification," "elimination," or "inconclusive" when compared to components from a particular firearm. *Id.* at 517–18. The court in *Boss* concluded "that the AFTE theory of identification … was sufficiently reliable under section 490.065.2." *Id.* at 518. The court in *Boss* also then noted that "[t]his conclusion is supported by the considerable weight of authority, which has

held firearm and tool mark identification admissible under the *Daubert* standards of admissibility." *Id.* (citing U*nited States v. Otero*, 849 F. Supp.2d 425, 437-48 (D.N.J. 2012)).

The Eastern District of our Court of Appeals reached the same conclusion in *State v. Mills*, 623 S.W.3d 717 (Mo. App. E.D. 2021). *Mills* followed *Boss* in recognizing that "toolmark examination evidence was sufficiently reliable, even if results somewhat rely on a 'subjective analysis' and the examiner's expertise and experience." *Id.* at 732 (quoting *Boss*, 577 S.W.3d at 518). The court in *Mills* even concluded that the trial court did not abuse its discretion by denying the defendant's request for a *Daubert* hearing before allowing the toolmark examiner to testify: "A *Daubert* hearing was not required to admit Menendez's testimony, which we find sufficiently reliable." *Id.*

At the *Daubert* hearing held here, Defendant's counsel conceded an absence of Missouri appellate decisions excluding firearm and toolmark comparisons. He nonetheless urges this Court to exclude such evidence here, claiming that the courts in *Boss* and *Mills* simply did not have the benefit of Dr. Scurich's testimony, and he points this Court to cases from other jurisdictions which have considered his opinions and criticisms of studies supporting the reliability of toolmark-comparison evidence.

One such opinion of Dr. Scurich is his criticism of what can be termed "set-to-set" studies. In a set-to-set study, examiners are given a set of unknown ammunition components as well as a set of known origin and tasked with matching the unknown with the known. But by being grouped together in a set, each identification reduces the number of remaining options, making each successive pair easier to match. With what Dr. Scurich characterized as a Sudoku-like solvability, set-to-set studies

are of little use for determining actual false-positive rates. He suggests this would explain why (for example) several studies report an error rate of zero. One such study was referenced in *Boss*, where the State's expert "testified regarding a study that involved 7,500 comparisons conducted by more than 500 examiners throughout 20 countries, which did not produce a single false-positive identification." *Boss*, 577 S.W.3d at 518. While the level of consistency in such studies still belies Defendant's assertion that toolmark examination should be merely "characterized as a subjective art," (Def. Motion ¶ 4), Dr. Scurich's critique of set-to-set studies as a reliable source of known or probable error rates appears to be well-founded.

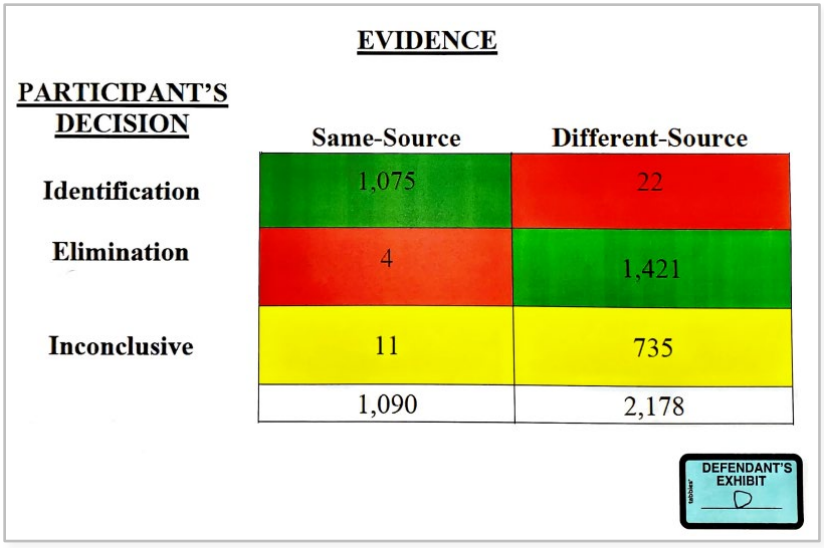
Even setting aside the set-to-set studies, however, there are sample-to-sample or "pairwise" studies with consistently low false-positive rates. The two primarily discussed at the hearing were referred to as the Ames I[1] and Ames II[2] studies. Ames I had a false-positive rate of about 1.1%. It involved 218 examiners evaluating 3,268 pairs of cartridge cases. 1,090 of those were same-source pairs, for which "identification" would be the correct answer. The examiners correctly identified 1,075 of the 1,090, with 11 marked "inconclusive" and just 4 marked "elimination." The remaining 2,178 were different-source pairs, for which "elimination" would be correct. The examiners correctly chose "elimination" for 1,421 of the 2,178, with 735 marked "inconclusive" and 22 marked "identification."

The subject examiners' decisions in Ames I are depicted in Defendant's Hearing Exhibit D, a summary chart which Dr. Scurich prepared from the original study data:

---

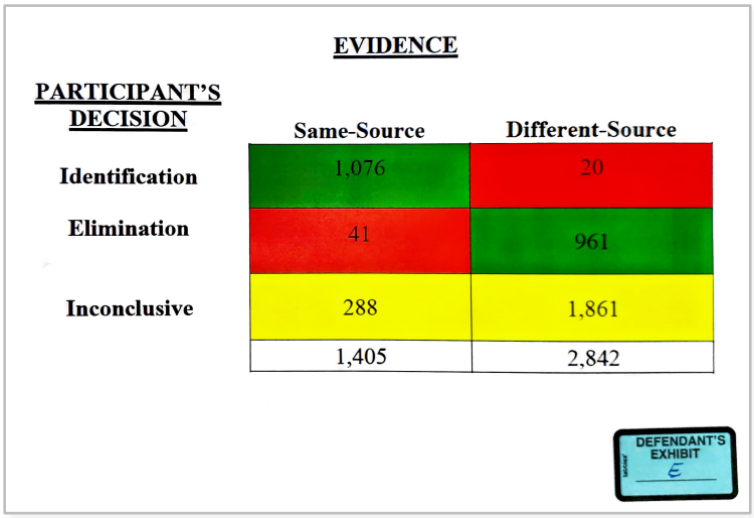
[1] David P. Baldwin et al., *A Study of False-Positive and False-Negative Error Rates in Cartridge Case Comparisons* (2014) (available at https://www.ojp.gov/pdffiles1/nij/249874.pdf).

[2] Stanley J. Bajic et al., *Report: Validation Study of the Accuracy, Repeatability, and Reproducibility of Firearm Comparisons* (Ames Laboratory-US DOE 2020) (Tech. Rep. #IS-5207).



PARTICIPANT'S DECISION	EVIDENCE: Same-Source	EVIDENCE: Different-Source
Identification	1,075	22
Elimination	4	1,421
Inconclusive	11	735
	1,090	2,178

Dr. Scurich acknowledged that in the entire Ames I study there were only 4 false-negative exclusions and 22 false-positive identifications. He argues, however, that counting the "inconclusives" in the overall number but not scoring them as errors gave examiners a "free pass" and artificially decreases the error rate. Taken to its logical extreme, this approach would allow examiners to mark every pair as "inconclusive" and still have an error rate of zero. He has similar criticisms for the Ames II study, which is depicted in Defendant's Hearing Exhibit E:



PARTICIPANT'S DECISION	EVIDENCE: Same-Source	EVIDENCE: Different-Source
Identification	1,076	20
Elimination	41	961
Inconclusive	288	1,861
	1,405	2,842

The Ames II study looked at both expended bullets and spent cartridge casings. The reported false positive rate is 0.656% for bullets and 0.933% for cartridge cases. But as noted by Dr. Scurich, the calculation of these rates—as with Ames I—include the "inconclusives" as correct answers, contrary to the guidance of the 2016 President's Council of Advisors on Science and Technology Report ("PCAST Report"). To correct for this, the "inconclusives" would need be removed from the overall number being counted—in other words, taken out of the denominator. If that were done, in Ames I the error rate would increase to around 1.5%; in Ames II, it would increase to around 2% for bullets and 1.86% for cartridge cases.

Dr. Scurich suggests an even more extreme approach, however. Rather than exclude the "inconclusives," he suggests an alternate scenario of counting them as errors on par with false-positives or false-negatives. In other words, rather than taking them out of the denominator, adding "inconclusives" to the numerator. For Ames I, this alternate approach would inflate the error rate to 34.7%. For Ames II, depending on which gradations of "inconclusive" designations were treated as errors, the error rates would be between 23% and 54% for the bullets, and 12.9% to 37.9% for the cartridge cases.

The approach of setting aside the "inconclusive" decisions has received some support as better aligning with real-world use of toolmark-comparison evidence "because evidence used against a defendant will typically be based on conclusive, rather than inconclusive, examinations." *United States v. Felix*, 2022 WL 17250458, at *16 (D.V.I. Nov. 28, 2022) (quoting PCAST Report at 153). But Dr. Scurich's alternative approach of equating an "inconclusive" with a false-positive error has been roundly rejected by those same courts, as it seeks to replace an artificially *deflated*

error rate with an artificially *inflated* error rate. *Id.* at *16. Even in *Tibbs*—the case that otherwise appears to be the most receptive to Dr. Scurich's criticisms generally—the court still rejected as illogical the notion that "inconclusive" responses should be classed with other errors:

> Dr. Scurich opines, based on principles of mathematics and statistics in particular, that such responses should be viewed as false positive errors (i.e., included among false identifications), but such a characterization fails to make logical sense: while under laboratory conditions such inconclusives are surely some type of error, it does not follow that inconclusives are functionally the same as a false conclusion by an examiner who attributes a cartridge casing to a gun that did not fire it.

U.S. v. Tibbs, 2019 WL 4359486, at *17 (D.C. Super. Sep. 05, 2019). Other decisions have been similarly unreceptive to this theory:

> With respect to … Dr. Scurich's opinions that the error rates are unreliable because they do not include inconclusive results, this Court finds those opinions unsupported. Or, perhaps more accurately, the Court finds that while the relatively high rate of inconclusive results may be relevant to certain policy determinations—such as failing to eliminate a suspect from consideration—those concerns do not relate to the issue of paramount importance to trial courts; i.e., the false positive rate which may result in a wrongful conviction.

United States v. Rhodes, 2023 WL 196174, at *4 (D. Or. Jan. 17, 2023). The consensus view, with which this Court agrees, is that Dr. Scurich may well be correct that "an inconclusive result is an error insofar as it means the methodology did not produce an answer," but that does not make it "an error in the sense that it falsely attributes a cartridge or casing to the

wrong firearm." *State v. Raynor*, 2024 WL 3579515, at *9 (Conn. Super. Ct. Apr. 16, 2024) (quoting *Rhodes*, 2023 WL 196174, at *4).

Even accepting the higher error rates that would result from disregarding the "inconclusives," the Ames I and Ames II studies suggest that when examiners reached decisions of "identification" or "exclusion," their decisions were correct around 98% of the time. As several courts have noted, this is well within the range suggested by PCAST as an acceptable error rate: "The 2016 PCAST Report, on which Defendant primarily relies to critique the majority of the validity studies, posits that an acceptable error rate from a scientific perspective is 5%. Thus a 2.2% error rate would still be significantly lower than the recommended threshold by scientific experts." *United States v. Chavez*, 2021 WL 5882466, at *4 (N.D. Cal. Dec. 13, 2021); *see also Rhodes*, 2023 WL 196174, at *4. This Court agrees and concludes that firearm and toolmark identification is reliable enough to allow expert testimony on the subject under Section 490.065.2.

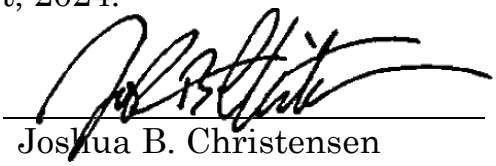
To be sure, the rate of false positives in studies do not automatically translate to the number of false positives in real-life casework, much less to any one case. As Dr. Scurich observed, the behavior of examiners as subjects in a study may be affected by the simple knowledge that they are being studied. Examiners in a study may also be impacted by features of a study's design that might encourage "inconclusive" answers.[3] The

[3] As noted by the State, differences between studies and real-world conditions may well cut the other direction too. For example, the potential consequences of false identifications or exclusions in actual cases arguably creates a strong incentive to favor the "inconclusive" option in actual casework. Additionally, the error rates from studies capture only an examiner's individual conclusions at a given moment, divorced from any quality-control measures that examiner would employ in real casework. In accredited labs such as the Missouri State Highway Patrol Crime Laboratory, examiners have the benefit of "a second examiner who independently examines the results of the first examiner." *Chavez*, 2021 WL 5882466, at *4. And at trial, of course, there is opportunity for the examiner—and the jury—to evaluate the effectiveness of any cross-examination or contrary expert opinions.

certainty of any particular identifications may also be questioned on any number of other grounds: concerns about repeatability or reproducibility, the particular manner of any individual examination, or outside information given to an examiner beyond the materials given for comparison, to name just a few. And even a 2% error rate, while reliable enough to be admissible, is not insignificant given the stakes in criminal trials.

But such disputes about the certainty of an expert's opinions "should be tested by the adversary process with competing expert testimony and cross-examination, rather than excluded by the court at the outset." *Johnson*, 754 F.3d at 562 (citing *Daubert*, 509 U.S. at 590). In allowing for the admission of toolmark examination evidence under Section 490.065, *Boss* specifically emphasized that the defendant was free to challenge the expert's conclusions and point out the weaknesses of his analysis to the jury during cross-examination. *Boss*, 577 S.W.3d at 519. "Vigorous cross-examination, presentation of contrary evidence, and careful instruction on the burden of proof are the traditional and appropriate means of attacking shaky but admissible evidence." *Id.* (quoting *United States v. Davis*, 103 F.3d 660, 674 (8th Cir. 1996)). This Court is bound to follow *Boss* and *Mills*. Because the State's proposed testimony meets the standards of Section 490.065.2, and the State has agreed to limit testimony on the topic to the DOJ standards to the greatest extent possible (a topic the parties should be prepared to address at the pretrial conference), Defendant's request to exclude expert testimony is respectfully denied.

SO ORDERED on this 16th of August, 2024.



Joshua B. Christensen
Circuit Judge, Division 6