

RECOMMENDED FOR PUBLICATION
Pursuant to Sixth Circuit I.O.P. 32.1(b)

File Name: 21a0057p.06

UNITED STATES COURT OF APPEALS

FOR THE SIXTH CIRCUIT

UNITED STATES OF AMERICA,

Plaintiff-Appellant,

v.

DANIEL GISSANTANER,

Defendant-Appellee.

}
} No. 19-2305
}

Appeal from the United States District Court
for the Western District of Michigan at Grand Rapids.
No. 1:17-cr-00130-1—Janet T. Neff, District Judge.

Argued: January 29, 2021

Decided and Filed: March 5, 2021

Before: SUTTON, BUSH, and MURPHY, Circuit Judges.

COUNSEL

ARGUED: Justin M. Present, UNITED STATES ATTORNEY'S OFFICE, Grand Rapids, Michigan, for Appellant. Joanna C. Kloet, OFFICE OF THE FEDERAL PUBLIC DEFENDER, Grand Rapids, Michigan, for Appellee. **ON BRIEF:** Justin M. Present, UNITED STATES ATTORNEY'S OFFICE, Grand Rapids, Michigan, for Appellant. Joanna C. Kloet, OFFICE OF THE FEDERAL PUBLIC DEFENDER, Grand Rapids, Michigan, for Appellee. Maneka Sinha, UNIVERSITY OF MARYLAND, Baltimore, Maryland, M. Katherine Philpott, VIRGINIA COMMONWEALTH UNIVERSITY, Richmond, Virginia, for Amici Curiae.

OPINION

SUTTON, Circuit Judge. At issue in this case is the reliability of a form of DNA-sorting evidence under Rule 702 of the Federal Rules of Evidence and *Daubert v. Merrell Dow Pharmaceuticals, Inc.*, 509 U.S. 579 (1993).

I.

Daniel Gissantaner became embroiled in an argument with his neighbors. One neighbor called 911, telling the dispatcher that Gissantaner, a convicted felon, had a gun. The police responded to the call and found a pistol inside a chest in Gissantaner's house. The chest belonged to Gissantaner's roommate. When the government charged Gissantaner with possessing a firearm as a felon, it used DNA-sorting evidence, called STRmix, to link Gissantaner to the gun.

Gissantaner moved to exclude the evidence as unreliable under Evidence Rule 702. *See Daubert v. Merrell Dow Pharms., Inc.*, 509 U.S. 579 (1993). Gissantaner and the government retained experts, who took competing positions on the issue. The district court in turn appointed two experts of its own to consider the question. One of these experts said that STRmix evidence is reliable in general and as applied to this case; the other said it is reliable in general but not as applied to this case. The court excluded the STRmix evidence as unreliable. The government filed this interlocutory appeal. *See* 18 U.S.C. § 3731.

II.

DNA evidence has transformed criminal investigations, trials, and post-conviction proceedings. Since the late 1980s, it has become a staple of law-enforcement investigations, whether to track down the guilty or to liberate the innocent. *See* Robert J. Norris, *Exonerated* 35 (2017). In contrast to blood types and enzymes, an individual's DNA is unique. *United States v. Bonds*, 12 F.3d 540, 550 (6th Cir. 1993). No one else shares it, save an identical twin. That

makes DNA evidence highly pertinent to the work of forensic scientists tasked with investigating crimes.

Complications arise when a sample of evidence, usually a batch of cells included in fluid emitted by an individual or left on an item by touch, contains the DNA of more than one person. Consider this case. Police officers collected the relevant evidence on a gun found in a chest owned by Gissantaner's roommate. They collected the "touch DNA"—skin cells found on the gun—and submitted it for analysis at Michigan's statewide law-enforcement laboratory. Based on the genetic material in the mixture, an analyst determined that the DNA recovered from the weapon came from three people.

In the early years of working with DNA to solve crimes, forensic scientists used one technique—visual comparison—to handle DNA samples with single and multiple strains. While scientists eventually became adept at handling samples with up to two strains of DNA, they faced difficulties beyond that.

A digression into how forensic scientists use DNA evidence helps to explain why. Some spots on the human genome, named loci, contain segments of DNA code that vary widely from one person to another. Each variation is called an allele, and a person generally has two alleles at each locus. Because the alleles vary, no two people are likely to have matching alleles at any given locus. A greater number of matching alleles at a greater number of loci increases the probability that the person of interest matches the sample. *See* Special Master's Report at 17–24, *United States v. Lewis*, 442 F. Supp. 3d 1122 (D. Minn. 2020) (No. 18-194).

One challenge with using mixtures involving several people is that each person might have contributed zero, one, or two alleles at each locus. (Although people have two alleles at each locus, one or more of the alleles might not have made it into the mixture.) If a mixture contains five different alleles at one locus, that could suggest it involves at least three people, with two contributors donating two alleles and a third contributor donating one. But other possibilities remain. It could be that one contributor gave two alleles and three other contributors gave one allele at the locus, suggesting a four-person mixture.

Even with these variations, visual examinations of the alleles in a mixture still allow examiners to estimate how many DNA sources a touch or fluid sample contains. That is just what forensic scientists did in the first decades after they began using DNA to solve crimes.

Visual examinations come with subjective risks, however. The inspection might over-count or under-count the percentage of each individual's contribution to the sample, to the extent inspectors calculated the percentage at all, or might mistake the number of people who contributed to it. *Cf. id.* at 21, 23–24. Visual inspection runs the risk of cognitive biases, too. Studies suggest that an examiner's knowledge of the case—other evidence about the suspect—affects interpretations, frequently not in the suspect's favor. On top of all that, the calculations used to determine the probability of the mixture's occurring by chance, as opposed to coming from the suspect, have to be simplified because human beings, Alan Turing included, are not computers.

Enter STRmix and other DNA-sorting products. Starting in the late 1990s, forensic scientists innovated products to improve investigations of multi-person DNA samples. The idea is to combine the tools of DNA science, statistics, and computer programming to mitigate the risks from subjective assessments of multi-person DNA samples. The software in the end helps to measure the probability that a mixture of DNA includes a given individual's DNA.

In addition to mitigating the risks of human error, the software processes more potential interpretations of a mixture in less time. If an analyst remains unsure whether a sample contains the DNA of three persons or four, she can use the software to crunch the numbers quickly in both ways. The software also mitigates the effect of cognitive bias, as the software does not know the other facts of the case. *Id.* at 23–27. While the software does not eliminate the ever-present risks of human error, it “clearly represent[s] a major improvement over purely subjective interpretation.” R.41-17 at 93.

Forensic labs today use several probabilistic genotyping software programs, including STRmix, LRmix, Lab Retriever, likeLTD, FST, Armed Xpert, TrueAllele, and DNA View Mixture Solution. The product used in this case, STRmix, was created in 2011 by John

Buckleton, a civil servant and forensic scientist who works for the Institute of Environmental Science and Research, a government agency in New Zealand.

Roughly 200 forensic science laboratories exist in the United States. Most are affiliated with a government. Michigan, for example, has the Michigan State Police laboratory, dedicated to providing scientific and technological support to law enforcement throughout the State. Over 45 of these law-enforcement laboratories use STRmix, with more on the way. R.139 at 5–6 (noting that 68 laboratories are in the process of validating STRmix). About ten other laboratories use similar products. The FBI uses STRmix. A license for unlimited use of STRmix costs about \$27,000, with proceeds supporting the work of the New Zealand Institute.

The Biology DNA Unit of the Michigan State Police laboratory has used STRmix for six years. The laboratory received training on the software starting in March 2015, and it began using the software for cases in February 2016, about three and a half months before receiving the sample in this investigation.

In Gissantaner’s case, an analyst with the Michigan State Police laboratory took information about the DNA present in the mixture and entered it into STRmix to estimate how much of the DNA came from each person. The resulting DNA profile summary said that one person “contributed” 68% of the DNA in the mixture, a second contributed 25%, and a third contributed 7%. The third contributor supplied 8 or 9 cells (approximately 49 picograms) to the mixture.

STRmix compared the DNA of the suspect—Gissantaner—to this profile with the goal of ascertaining a “likelihood ratio” about his potential contribution to the sample. R.77 at 34. A comparison of Gissantaner’s DNA to the profile suggested that he matched the third contributor, generating a likelihood ratio of 49 million to 1. More precisely, if less accessibly, that means the DNA “profile is 49 million times more likely if [Gissantaner] is a donor than if he is not.” *Id.* at 48.

The two “ifs” capture a qualification. The likelihood ratio tells us only that, in the abstract and without considering any other evidence in this case, it would be unusual if this DNA contained no DNA contributed from Gissantaner. The ratio does not on its own tell us how

likely it is that Gissantaner illegally possessed a firearm. Determining whether it is likely that Gissantaner, as opposed to someone else, contributed to the mixture requires looking at other facts beyond the scope of DNA evidence. Perhaps other people with similar profiles, say relatives of Gissantaner, were nearby. Perhaps the roommate had locked the chest and lost the key long before Gissantaner moved in and had a chance to touch the gun. Or perhaps the DNA landed on the gun without Gissantaner touching it, say when he and his roommate shook hands or when they each touched the same object before his roommate subsequently touched the gun. Even with these qualifications, the evidence remains significant. By accepted measures in the forensic community, a profile match of 1 in 49 million amounts to “very strong support” that Gissantaner somehow contributed DNA to the mixture. R.41-10 at 3.

These two bottom-line conclusions bring into view the need to use evidence of this sort carefully. Such conclusions—the 49-million-to-1 ratio and “very strong support”—can be highly probative of guilt or innocence. Yet the mechanisms for obtaining them—the software, the science—are beyond the ken of most jurors and judges. If highly consequential evidence emerges from what looks like an indecipherable computer program to most non-scientists, non-statisticians, and non-programmers, it is imperative that qualified individuals explain how the program works and ensure that it produces reliable information about the case. All of this explains why the courts have developed reliability standards for admitting evidence of this type and why a functioning adversarial system remains critical in handling it.

III.

Rule 702 of the Federal Rules of Evidence sets the framework for determining whether to admit scientific and other technical evidence in federal civil and criminal cases. It permits an expert to testify about scientific knowledge 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.” Fed. R. Evid. 702. The same set of questions applies to expert testimony and science-based test results. *Compare Johnson v. Manitowoc Boom Trucks, Inc.*,

484 F.3d 426, 427–29 (6th Cir. 2007), with *United States v. Semrau*, 693 F.3d 510, 516, 520 (6th Cir. 2012).

Four inquiries guide the reliability analysis: Is the technique testable? Has it been subjected to peer review? What is the error rate and are there standards for lowering it? Is the technique generally accepted in the relevant scientific community? *Daubert*, 509 U.S. at 593–94. Multi-factor tests, especially non-exhaustive tests, run the risk of obscuring the core inquiry. The key handholds of Rule 702 thus bear repeating: To be admissible, any relevant scientific or technical evidence must be the “product of reliable principles and methods” and must have been “reliably applied” in the case. That is what matters most. Otherwise, the central point of *Daubert*, to establish that Evidence Rule 702 “displaced” the common law *Frye* test, would be lost and would lead to the replacement of an old common law test with a new (harder to pronounce) common law test. Compare *Frye v. United States*, 293 F. 1013, 1014 (D.C. Cir. 1923), with *Daubert*, 509 U.S. at 585–89. That is not progress.

Through it all, the district court has a “gatekeeping role” in screening expert testimony to ensure that only reliable testimony and evidence go to the jury. *Daubert*, 509 U.S. at 597. We give fresh review to a district court’s framing of the legal standard, *United States v. Pugh*, 405 F.3d 390, 397 (6th Cir. 2005), and abuse-of-discretion review to its admissibility decision, *Kumho Tire Co. v. Carmichael*, 526 U.S. 137, 152 (1999).

IV.

Measured by Evidence Rule 702 and a proper framing of the *Daubert* factors, the DNA evidence should be admitted on this record.

Testability. An untestable scientific theory is all theory and no science. In the absence of proof that a technology “can be . . . tested,” *Daubert*, 509 U.S. at 593, there is no way to show whether it works (its “refutability” or “falsifiability,” a scientist would say) and no way to give it “scientific status.” *Id.*; *United States v. Bonds*, 12 F.3d 540, 559 (6th Cir. 1993). The question on the table is whether a method can be “assessed for reliability,” not whether it always gets it right. Fed. R. Evid. 702 advisory committee’s note to 2000 amendment; *Bonds*, 12 F.3d at 558–59. Disputes about the “adequacy of the [theory’s] testing” or about the “accuracy of

[a theory's] results," generally speaking, provide grist for adversarial examination, not grounds for exclusion. *Bonds*, 12 F.3d at 558–59; see *United States v. Mitchell*, 365 F.3d 215, 238 (3d Cir. 2004); see also *City of Pomona v. SQM N. Am. Corp.*, 750 F.3d 1036, 1046 (9th Cir. 2014); *United States v. Baines*, 573 F.3d 979, 989–90 (10th Cir. 2009).

STRmix can be tested. Using “lab-created mixtures,” in which the actual contributors of the DNA samples are known, scientists have tested STRmix to gauge the reliability of the technology. R.146-14 at 2. Suppose that one person, Aaron, contributed to a lab-created mixture, but another, Britney, did not. Forensic scientists can test STRmix to see whether it suggests that Aaron is a match for the mixture, but Britney is not. If STRmix suggests that Aaron is not a match for the mixture (by outputting a low likelihood ratio), that would be a false negative. If STRmix suggests that Britney is a match for the mixture (by outputting a high likelihood ratio), that would be a false positive. Each possibility shows that STRmix is testable, that lab-created mixtures offer a way to “assess[] [the] reliability” of STRmix. Fed. R. Evid. 702 advisory committee’s note to 2000 amendment.

The record from the evidentiary hearings in this case provides a long proof that STRmix is testable and refutable. Almost all of the evidence in the hearings went to these points: How often is it accurate? How often is it not? Similar evidence of DNA testability has sufficed before in our circuit. *Bonds*, 12 F.3d at 558.

(The astute reader might remind us that *Daubert* asks whether the scientific theory at issue “can be (and has been) tested.” 509 U.S. at 593. It is not clear what the parenthetical means. The rest of the *Daubert* considerations—including the next one, peer review—all turn in one way or another on what actual testing of the theory reveals in terms of reliability. At all events, the point makes no difference here. STRmix indeed “(has been) tested” many times before, as the rest of this opinion confirms.)

Peer review. Subjecting a new technology to “peer review and publication” offers another measure of reliability. *Id.* The “key” is whether “the theory and procedures have been submitted to the scrutiny of the scientific community.” *Bonds*, 12 F.3d at 559. Publication in a peer-reviewed journal typically satisfies this consideration. See *Daubert*, 509 U.S. at 594.

Peer review is not student review. It “conditions publication on a bona fide process” of review by other scientists and experts in the field. *Daubert v. Merrell Dow Pharms., Inc.*, 43 F.3d 1311, 1318 n.6 (9th Cir. 1995). The scientific community uses different conventions for publication from most journals published in the legal community. No offense to former, current, and future members of law journals everywhere: But it is one thing to convince lawyers in training to publish a piece; it is quite another to convince peers in a professional community to publish a piece. That is why readership and citation are pivotal when it comes to legal scholarship and why publication itself is noteworthy in scientific scholarship—and ultimately why publication in a peer-reviewed journal alone typically satisfies this *Daubert* inquiry. See *Ruiz-Troche v. Pepsi Cola of P.R. Bottling Co.*, 161 F.3d 77, 84–85 (1st Cir. 1998); *In re Paoli R.R. Yard PCB Litig.*, 35 F.3d 717, 781 (3d Cir. 1994); *United States v. Brown*, 973 F.3d 667, 704 (7th Cir. 2020); cf. *Allison v. McGhan Med. Corp.*, 184 F.3d 1300, 1316 (11th Cir. 1999); *Summit 6, LLC v. Samsung Elecs. Co.*, 802 F.3d 1283, 1295 (Fed. Cir. 2015).

For like reasons, this factor does not demand independent authorship—studies done by individuals unaffiliated with the developers of the technology. Independent studies, to be sure, advance the cause of reliability. *Bonds*, 12 F.3d at 560. But they are not indispensable. Peer review contains its own independence, as it involves “anonymously reviewing a given experimenter’s methods, data, and conclusions on paper.” *Mitchell*, 365 F.3d at 238. If experts “have other scientists review their work” and if the other scientists have the chance to identify any methodological flaws, that usually suffices. *Mitchell v. Gencorp Inc.*, 165 F.3d 778, 784 (10th Cir. 1999). When scientific research is accepted for publication by a reputable journal following the “usual rigors of peer review,” that represents “a significant indication that it is taken seriously by other scientists, i.e., that it meets at least the minimal criteria of good science.” *Daubert*, 43 F.3d at 1318.

STRmix clears this bar. At the time of the *Daubert* hearing in the district court, more than 50 published peer-reviewed articles had addressed STRmix. According to one expert, STRmix is the “most tested and most . . . peer reviewed” probabilistic genotyping software available. R.77 at 82. At least two of the studies were done by individuals unconnected to the

development of the software. This plainly suffices. *Bonds*, 12 F.3d at 559–60; *cf. Gross v. Comm’r*, 272 F.3d 333, 340–41 (6th Cir. 2001).

Error rate and standards to lower it. Even ten witnesses to a crime come with risks of error. So too for DNA evidence. This consideration looks to the error rate of the technology and to whether the scientific community has established standards that forensic scientists can use to mitigate the risk of error. *Daubert*, 509 U.S. at 594.

Think about Gissantaner’s case to see the point. The government would like to use STRmix to match Gissantaner to the DNA on a gun. That is not a good idea—not the “product of reliable principles and methods” under Rule 702—if STRmix has a high error rate, if it has trouble “avoid[ing]” “false positives,” and if there are no standards or guidelines to avoid or lessen these risks. *Bonds*, 12 F.3d at 559 (quotation omitted); *Mitchell*, 365 F.3d at 241.

How often, then, does STRmix falsely suggest a suspect matches a DNA sample? Not often, the evidence suggests. When examining “false inclusions,” one peer-reviewed study concluded, based on an analysis of the DNA of 300,000 people who were known not to be in a mixture, that STRmix had accurately excluded the non-contributors 99.1% of the time. Just 1% of the time, in other words, it gave a likelihood ratio suggesting that someone was included in the mixture who was not actually included in it. Most of these false inclusions, moreover, were associated with low likelihood ratios—meaning that, under STRmix’s own estimates, the confidence that the person was included was low. A likelihood ratio of 100 to 1 is more likely to produce a false inclusion than a likelihood ratio of 1 million to 1. In this instance, the likelihood ratio was 49 million to 1.

One explanation for the low error rate is the existence of standards to guide the use of STRmix and other probabilistic genotyping software, for the two are “[c]losely related.” *Mitchell*, 365 F.3d at 241. The Scientific Working Group on DNA Analysis Methods, a national association of forensic laboratories sponsored by the FBI, has produced guidelines governing the use of this kind of software, guidelines that the Michigan State Police laboratory used in this case.

General acceptance in the scientific community. One might still be skeptical of the reliability of a relatively new technology like STRmix, and rightly so, if the relevant scientific community has not yet accepted its use. *Daubert*, 509 U.S. at 594. The question for debate is “general acceptance,” not uniform acceptance within the community. *See Bonds*, 12 F.3d at 562. Nor must the science be beyond reproach. *Id.* What matters is whether the relevant scientific community accepts the software. *See Daubert*, 509 U.S. at 594; *see also Wilden v. Laury Transp., LLC*, 901 F.3d 644, 654–55 (6th Cir. 2018). After that, the long-tested cauldron of cross-examination, not exclusion, is the place to go for accuracy. “[C]onventional [trial] devices, rather than wholesale exclusion under an uncompromising ‘general acceptance’ test, are the appropriate safeguards where the basis of scientific testimony meets the standards of Rule 702.” *Daubert*, 509 U.S. at 596. For a technology that is widely used, controversies over its use in a given case usually will be left to the jury. *See United States v. Jones*, 965 F.3d 149, 160 (2d Cir. 2020).

STRmix satisfies this consideration. It has garnered wide use in forensic laboratories across the country. More than 45 laboratories use it, including the FBI and many state law enforcement agencies. At this point, STRmix is the “market leader in probabilistic genotyping software.” R.146-1 at 17.

Consistent with this reality, numerous courts have admitted STRmix over challenges to its general acceptance in the relevant scientific community. *See United States v. Lewis*, 442 F. Supp. 3d 1122, 1155 (D. Minn. 2020) (“[T]here is no doubt that STRmix has gained general acceptance.”); *United States v. Washington*, No. 8:19CR299, 2020 WL 3265142, at *2 (D. Neb. June 16, 2020) (“Authority and evidence demonstrate that STRmix is generally accepted by the relevant community.”); *People v. Blash*, No. ST-2015-CR-0000156, 2018 WL 4062322, at *6 (V.I. Super. Ct. Aug. 24, 2018); *People v. Muhammad*, 931 N.W.2d 20, 30 (Mich. Ct. App. 2018); *People v. Bullard-Daniel*, 42 N.Y.S.3d 714, 724–25 (N.Y. Co. Ct. 2016); *United States v. Christensen*, No. 17-CR-20037-JES-JEH, 2019 WL 651500, at *2 (C.D. Ill. Feb. 15, 2019) (“STRmix has been repeatedly tested and widely accepted by the scientific community.”); *United States v. Oldman*, No.18-CR-0020-SWS, ECF No. 227 at *16 & n.5 (D. Wyo. Dec. 31, 2018) (collecting cases); *United States v. Russell*, No. CR-14-2563 MCA, 2018 WL 7286831, at

*7–8 (D.N.M. Jan. 10, 2018) (“[STRmix’s] analyses are based on calculations recognized as reliable in the field.”); *United States v. Pettway*, No. 12-CR-103S (1), (2), 2016 WL 6134493, at *1 (W.D.N.Y. Oct. 21, 2016) (discussing “exhaustive[] research[]” concluding that “the scientific foundations of the STRmix process are based on principles widely accepted in the scientific and forensic science communities”). The Second Circuit determined that the scientific community accepted a different (but similar) DNA-sorting software, Forensic Statistical Tool, even though just one laboratory had used it. *Jones*, 965 F.3d at 156, 162.

General acceptance of probabilistic genotyping software, moreover, has led to its use in inculpatory and exculpatory settings alike. See Erik Ortiz, A Texas jury found him guilty of murder. A computer algorithm proved his innocence., <https://news.yahoo.com/prison-murder-computer-algorithm-helped-105609137.html> (last visited March 3, 2021); Jason Hanna & Nick Valencia, Thanks to a new DNA analysis, a Georgia man is exonerated of rape and freed from prison after 17 years, <https://www.cnn.com/2020/01/10/us/georgia-kerry-robinson-released/index.html> (last visited March 3, 2021).

All in all, STRmix satisfies Rule 702 and the case law construing it. In the words of Rule 702, it is the “product of reliable principles and methods.”

V.

But were those principles “reliably applied” in this case, as Rule 702 also asks? Yes, as we explain and as the record in this case demonstrates.

The Michigan State Police laboratory complies with the guidelines promulgated by the Scientific Working Group, as confirmed through an audit performed by the FBI. The forensic scientist who ran the sample in this case began training with STRmix more than a year before analyzing the sample. During that time, the laboratory tested its copy of the software, using lab-created mixtures to establish “internal validation” that STRmix reliably assisted the laboratory’s work. R.77 at 52. In addition to these artificial mixtures, the laboratory tested adjudicated-case samples, real samples from real crime scenes. It produced a summary explaining the results of its tests and offered additional data to supplement the summary. In

general, the laboratory's use of STRmix produced "significant likelihood ratios" for known contributors and improbable ones for known non-contributors. R.41-14 at 43.

The Michigan State Police laboratory's internal validation also included samples like the one in this case in which a minor contributor donated a small amount of DNA. It tested a mixture in which one contributor gave just 4% of the DNA (less than the 7% here) and another mixture in which the minor contributor gave only 26 picograms of DNA (less than the 49 picograms here). The laboratory also produced supplemental data showing that its internal validation included a lab-created mixture of 3.2% *and* 32 picograms and an adjudicated-case mixture of 4% *and* 10 picograms. The government offered to provide still more data for those interested, but the district court declined the offer.

STRmix also accounts for small amounts of DNA when it creates profile summaries. Because less DNA in a sample creates more uncertainty, STRmix generates lower likelihood ratios for low-quantity DNA mixtures than it otherwise would. The software also errs in the direction of the innocence of criminal suspects by making conservative estimates about the probability of a genetic pattern occurring.

The Michigan State Police laboratory has ample company in concluding that STRmix works at low levels of DNA. A peer-reviewed article compiling data from the internal validations of 31 independent laboratories indicated that STRmix had been validated with mixtures involving a minor contributor who supplied a small percentage of a mixture. The FBI's internal validation, also subjected to peer review, included mixtures in which the minor contributor contributed less than 7% and fewer than 49 picograms to the sample. Gissantaner did not introduce any studies showing STRmix is unreliable at low levels.

On this record, the "reliable principles and methods" underpinning STRmix were "reliably applied" in this case. Fed. R. Evid. 702. Any lingering concerns about how the Michigan State Police laboratory presented the information from its internal validation study or doubts about the reliability of STRmix at low levels of DNA can be hashed out through cross-examination or testing by the defendant. On appeal, Gissantaner and his team have not objected to the level of access the owners have provided to the program's source code or to their

ability to run their own tests at different parameters as well as to observe how the laboratory operates STRmix. While they have not used their access to this information yet to undermine the reliability of STRmix in general, they remain free to do so at trial.

VI.

Gissantaner resists this conclusion on several fronts, each unconvincing.

He claims that the standard of review favors him. That is true. But a district court may abuse its discretion by incorrectly framing the legal standard. *See United States v. Flowers*, 963 F.3d 492, 497 (6th Cir. 2020); *Pugh*, 405 F.3d at 397. The meaning of Evidence Rule 702, and for that matter the *Daubert* considerations, amounts to a legal question. *See United States v. Jones*, 107 F.3d 1147, 1154 (6th Cir. 1997); *Bureau v. State Farm Fire & Cas. Co.*, 129 F. App'x 972, 975 (6th Cir. 2005).

The district court framed several *Daubert* factors incorrectly. Start with testability. The district court pitched the question as “whether the use of STRmix has been adequately tested and validated, independently of the testing by the developer.” R.161 at 30. It then identified “shortcomings,” *id.* at 31, in the way that the Michigan State Police laboratory had displayed its test results, leading the court to conclude that the factor weighed “strongly against” admitting STRmix, *id.* at 35. Even “serious deficiencies” in testing, however, do not render a method untestable. *Bonds*, 12 F.3d at 559. At stake is “scientific validity,” not “scientific precision.” *Id.* at 558. Gissantaner’s, and the district court’s, “attempt[s] to refute the [government’s] theory and methods with evidence about deficiencies in both the results and the testing of the results,” amounts to a “conce[ssion] that the theory and methods can be tested.” *Id.* at 559. Although the independent experts in this case disagreed about the adequacy of the testing, that does not mean the theory is untestable or even that it has not been tested. *Id.*

Move to peer review. The district court’s position—that the factor requires studies authored or conducted independently of the developers of STRmix—misapprehends the inquiry. Independent authorship may (or may not) represent the scientific ideal, but submission to peer review generally suffices under *Daubert*. The court also overstated the risks of allowing the developers of a new technology to be the ones who test it. The key developer of STRmix is a

civil servant who works for New Zealand. Any revenue from sales of the software goes to a government agency, which by all appearances seems as focused on sparing the innocent as on convicting the guilty. What inculcates one day may exonerate the next with DNA-sorting evidence.

But even if that were not the case, even if STRmix had been developed by a for-profit entity, that would not belittle its reliability. Think of all the medical and scientific breakthroughs innovated by private corporations. Once one starts compiling that list, it is hard to stop. Come to think of it, how many of the vaccines for COVID-19 grew out of not-for-profit work? Peer review, together with the other *Daubert* factors, works well in handling the self-interest, whether modest or extravagant, that comes with any invention.

Close with general acceptance. In concluding that this “factor does not add weight for a finding that the STRmix DNA analysis is reliable,” R.161 at 43, the district court rooted its reasoning in the concern that STRmix “remains controversial” among a subset of the scientific community (computer scientists) and in cases involving small amounts of DNA. *Id.* But the existence of criticism, particularly as applied in specific cases, does not mean that STRmix has fallen short of “general” acceptance. The criticism at all events is overstated. Recall that the court appointed two experts under Evidence Rule 706. It then claimed that, because these two independent experts disagreed about using STRmix in this case, that meant it failed *Daubert*. But one of the experts, Dr. Krane, was not as independent as the court suggested. He was the president of the firm that employed Gissantaner’s expert, and he had previously worked on the case for Gissantaner, all points explained to the court by the government. That does not prohibit Dr. Krane from offering expert testimony about STRmix on behalf of Gissantaner. It just means he should not have been treated as an independent Rule 706 expert in the case. *See* 29 Charles Alan Wright & Victor Gold, *Federal Practice and Procedure* § 6304 (2d ed. 2016); *cf.* *Gates v. United States*, 707 F.2d 1141, 1144 (10th Cir. 1983) (per curiam). All told, there was no conflict between independent experts about the generally reliability of STRmix at low percentages and low weights. In truth, the only independent expert in the case found it reliable.

The district court’s misframing of these three factors provides one ground for our decision. Another is that the complete exclusion of such widely used scientific evidence, at least

on this record, would amount to an abuse of discretion anyway. The district court mainly questioned the reliability of this evidence because the Michigan State Police laboratory did not internally validate STRmix at low contribution and weight levels. As shown, however, the laboratory did validate STRmix at these levels, and so did the FBI. The district court's concerns with the laboratory's internal validation stemmed largely from the way the data was presented: (1) The validation summary did not mention mixtures similar to the one here—in which the minor contributor donated a small absolute amount of DNA (49 picograms) and a small percentage of the DNA in the mixture (7%)—and (2) the supplemental data, which explained that the Michigan State Police laboratory ran tests on similar mixtures, did not include the likelihood ratios or the false-positive rates from those tests.

Rule 702 does not require unstinting perfection in presenting test results. In light of the extensive evidence of STRmix reliability, including at low levels, these concerns were for the jury, not the court. *See Kumho Tire*, 526 U.S. at 153 (explaining that if reasonable experts disagree, that is usually for the jury, not the court). If the district court still wished to exclude the evidence based on a lack of data from the Michigan State Police laboratory's internal validation, it should have taken up the government's offer to present more evidence from the internal validation study. That is especially so given that the government had so far made good in responding to shifting requests from Dr. Krane throughout the case. Maybe the government has the precise false-positive rates from the laboratory's internal study; maybe it does not. If it does, its failure to anticipate the independent expert's exact request does not justify excluding STRmix. If it does not, any missing pieces can be used on cross-examination.

The court's reasoning in excluding the evidence gives pause in another respect. The key concern expressed by the court was the low percentage and low quantity (just 8 or 9 cells) supplied by the minor contributor. But it is the STRmix software, and nothing else, that supplied the factual premise for these points. It is puzzling to think that STRmix would be sufficiently reliable to show that the minor contributor gave just 7% of the sample but then to cast doubt on 7% samples in general and to do so when combined with low weight samples (49 picograms) in particular. Neither the district court nor Dr. Krane acknowledges, much less explains, the point.

As this last point suggests, the admissibility of this evidence is not the end of the road for Gissantaner. It may help him in some ways. Recall what the low 7% figure—created thanks to STRmix—means. It suggests that Gissantaner contributed only 8 or 9 cells to the mixture of DNA found on the gun. Holding a gun may not be the only way in which a small number of cells could land on a gun. Gissantaner was the roommate of the gun’s owner. Perhaps their interaction, their shared use of objects in the kitchen, or their shared touching of objects in the house affected the sample. That is what trials and juries are for. Both sides also will have ample opportunities to show whether the Michigan State Police laboratory’s testing of this sample was handled correctly and used correctly, and Gissantaner remains free to identify defects in the STRmix software. We have not hesitated to correct civil or criminal verdicts in which the trial record showed that the science was misused. *See Tamraz v. Lincoln Elec. Co.*, 620 F.3d 665 (6th Cir. 2010). Evidence Rule 403 also remains relevant to the inquiry. A district court concerned that the jury might misunderstand what the likelihood ratio means could require advocates to describe it in a way that will not generate “unfair prejudice” or “mislead[] the jury.” Fed. R. Evid. 403. Even with the admission of STRmix, the government still must show “beyond a reasonable doubt” that Gissantaner possessed the gun. *See Jackson v. Virginia*, 443 U.S. 307, 309 (1979). That may become a critical question if no other credible evidence shows that Gissantaner handled this gun.

We reverse.