Law, Probability and Risk (2013) **12**, 3–11 doi:10.1093/lpr/mgs031
Advance Access publication on November 5, 2012

Beyond uniqueness: the birthday paradox, source attribution and individualization in forensic science testimony

DAVID H. KAYE
Distinguished Professor and Weiss Family Scholar, School of Law and Graduate Faculty, Forensic Science Program, The Pennsylvania State University, University Park, Pennsylvania 16803, USA, kaye@alum.mit.edu

For many decades, forensic science identification experts have insisted that they can 'individualize' traces such as fingerprints and toolmarks to the one and only one object that produced them. They have relied on a theory of global uniqueness of patterns as the basis for such individualization. Although forensic practitioners and theorists are moving toward a more probabilistic understanding of pattern matching, textbooks and reference works continue to assert that uniqueness justifies individualization and that experience demonstrates discernible uniqueness. One response to the last claim applies a famous problem in probability theory—the Birthday Problem—to the forensic realm to show that even an extensive record of uniqueness does little to prove that all such patterns are unique. This essay describes the probabilistic reasoning and its limits. It argues that the logic of the Birthday Paradox does indeed undercut the theory of global, general uniqueness, but that the reasoning is logically compatible with opinion testimony that a specific object is nearly certain to be the source of a pattern or trace. It also notes some alternatives to categorical claims of individualization, whether those claims are based on the theory of global, general uniqueness or instead on some less sweeping and more defensible theory.

Keywords: uniqueness; individualization; source attribution; likelihood ratio; fingerprints; toolmarks.

1. Introduction

Forensic science is passing through an identity crisis. Every day, in every state, forensic analysts strive to establish the identity of impressions made by an object such as a gun barrel or the patterns deposited on a surface by fingers. Yet, some courts have begun to balk at these hoary claims of 'individualization'.[1] The classical theory is that for 'individual' as opposed to 'class' characteristics, 'individualization' is possible regardless of how many objects could be considered as possible sources.[2] A recent expression of this canon can be found in the *Fingerprint Sourcebook* published by the National Institute of Justice[3] and written and edited by the International Association for Identification (IAI)

[1] For example *United States* v. *Zajac*, No. 2:06-cr-00811-CW (D. Utah 16 September 2010) (limiting fingerprint testimony); *United States* v. *Glynn*, 578 F.Supp.2d 567, 574–75 (S.D.N.Y. 2008) (limiting firearms toolmark testimony).

[2] See, e.g. David H. Kaye, *Identification, Individuality, and Uniqueness: What's the Difference?*, 8 LAW, PROBABILITY & RISK 85 (2009).

[3] As 'the research, development, and evaluation agency of the U.S. Department of Justice,' NIJ's mission is to provide 'objective and independent knowledge and understanding of crime and justice issues through science'. NIJ, About NIJ, 4 April 2011, http://www.nij.gov/about/welcome.htm.

© The Author [2012]. Published by Oxford University Press. All rights reserved

Downloaded from http://lpr.oxfordjournals.org/ by guest on January 30, 2014

and the Scientific Working Group on Friction Ridge Analysis, Study, and Technology (SWGFAST). According to this reference work:

> If the rule of all pattern formations in nature being unique could definitely be demonstrated as false, or falsified, the rule would have to be altered. This falsification has never occurred. Based on observation, experimentation, and knowledge of pattern formations in nature (volar skin, other natural pattern formations, and their prints), the rule of law in forensic comparative sciences is: pattern formations in friction ridge skin cannot be replicated, and their prints can be individualized.[4]

The convenience of a scientific law that postulates that every natural pattern is uniquely identifiable cannot be gainsaid. The same law of nature also implies that given a discharged bullet recovered from a crime scene, firearms examiners can compare the impressions left on bullets fired from various guns. If they are sufficiently similar, the analyst can conclude—and later testify—that the bullet travelled through the barrel a specific gun. The classical theory of such individualization posits that: (1) every gun produces a unique pattern of striations on the bullets as they travel through the barrel, and (2) a skilled examiner can distinguish the pattern left by every gun from the pattern left by every other gun. Likewise, the classical theory of fingerprint identification demands that (1) every finger has a unique set of three-dimensional features that are reflected in two-dimensional fingerprint marks, and (2) a skilled analyst can distinguish the pattern left by every finger from the pattern left by every other finger. This is a theory of 'universal individualization' based on a premise of 'general uniqueness'.[5]

The problem is that the supposed law of nature is an article of faith rather than a scientifically supported conclusion.[6] The fact that the surfaces of every individual's fingers are unique (at a sufficient level of detail) is not, in itself, a persuasive argument for universal individualization.[7] The uniqueness argument proves too much, for just as every finger presumably is unique, so is *every* latent print, *even those from the same skin*.[8] Consequently,

> Uniqueness does not guarantee that prints from two different people are always sufficiently different that they cannot be confused, or that two impressions made by the same

[4] John R. Vanderkolk, *Examination Process*, *in* The Fingerprint Sourcebook Alan McRoberts ch. 9, 9–3 to 9–26 (U.S. Department of Justice, Office of Justice Programs, National Institute of Justice, Washington D.C., 2011).

[5] David H. Kaye, David E. Bernstein and Jennifer L. Mnookin, The New Wigmore, A Treatise on Evidence: Expert Evidence (2nd edn., Aspen Pub. Co., New York, NY, 2011). 'General uniqueness' means that every element of a set is distinguishable from every other element. 'Special uniqueness' means that a particular element is distinguishable from all others even if not all of the remaining elements are each distinguishable. Kaye, supra n 2.

[6] Compare Michael Saks and Jonathan J. Koehler, *The Individualization Fallacy in Forensic Science Evidence*, 61 Vand. L. Rev. 199 (2008) (arguing that nothing short of a census can prove uniqueness), with David H. Kaye, *Probability, Individualization, and Uniqueness in Forensic Science Evidence: Listening to the Academies*, 75 Brooklyn L. Rev. 1163 (2010) (arguing that proof short of a census can establish uniqueness in some populations, but agreeing that claims of global, general uniqueness are not supported). See also Simon A. Cole, *Forensics Without Uniqueness, Conclusions Without Individualization: The New Epistemology of Forensic Identification*, 8 Law, Probability & Risk 233 (2009); Jonathan J. Koehler and Michael J. Saks, *Individualization Claims in Forensic Science: Still Unwarranted*, 75 Brook. L. Rev. 1187 (2010) (modifying or 'refining' the claims in Saks & Koehler, supra).

[7] David Stoney, *Measurement of Fingerprint Individuality*, *in* Henry C. Lee and R. E. Gaensslen (eds), Advances in Fingerprint Technology, 331, p. 327–388, (2nd edn., CRC Press, Boca Raton, Florida, 2001).

[8] John Vanderkolk, Forensic Comparative Science: Qualitative Quantitative Source Determination of Unique Impressions, Images, and Objects 195 (Elsevier Academic Press, Burlington, Massachusetts, 2009); David Kaye, *Questioning a Courtroom Proof of the Uniqueness of Fingerprints*, 71 Int'l Statistical Rev. 521 (2003); cf. Simon Cole, Max Welling, Rachel Dioso-Villa and Robert Carpenter, *Beyond the Individuality of Fingerprints: A Measure of Simulated Computer Latent Print Source Attribution Accuracy*, 7 Law, Probability & Risk 165 (2008).

Downloaded from http://lpr.oxfordjournals.org/ by guest on January 30, 2014

> finger will also be sufficiently similar to be discerned as coming from the same source. The impression left by a given finger will differ every time, because of inevitable variations in pressure, which change the degree of contact between each part of the ridge structure and the impression medium.[9]

In short, the IAI-SWGFAST law of nature is inadequate to establish the theory of practical universal individualization. Nonetheless, it remains possible that all distinct impressions are distinguishable and that the failure to come across any distinguishable pair from the same source is powerful evidence of this state of affairs. In response to this possibility, Professor Jay Koehler offers a clever example to demonstrate the danger of inferring uniqueness even from a very large sample. He asks us to imagine that:

> [E]xactly 100 pairs of firearms out of an estimated 100,000 guns in a Texas town share indistinguishable gun barrel markings. If each of 100 firearms experts examined 10 pairs of guns from the town's gun population every day for 10 years (n=3,650,000 gun pairs), there is about a 93% chance that none of the indistinguishable pairs will have come under examination. That is, despite 1,000 "collective years" of forensic science experience (100 experts multiplied by 10 years), the failure to find even a single pair of guns with indistinguishable markings would offer little basis for drawing conclusions about whether gun barrel markings, even in this single town, are unique.[10]

This essay explores what the Texas Gun Case does—and does not—prove. Section 2 emphasizes that the hypothetical does indeed highlight the weakness of relying on the failure to falsify the postulated law of nature as proof of general (all-pairs) uniqueness in a large population of items. An expert relying on the argument for uniqueness in *The Fingerprint Sourcebook* to conclude that a bullet came from a matching gun in the Texas town therefore would be skating on thin ice. But Section 3 shows that a source attribution is very probably correct under the circumstances of the hypothetical. As a result, an expert testifying to an opinion that a bullet almost certainly came from a matching gun in the Texas town would be on *terra firma*. With these points established, Section 4 discusses reasonable limits on source-attribution testimony and a few alternatives to this form of presenting forensic identification evidence.

2. The Texas town and the birthday paradox

The Texas Gun Case is a variation on the famous Birthday Problem.[11] In its simplest form, the problem is this: Assuming that everyone is equally likely to be born on any day of the year (and that no one is ever born on February 29), how many people must enter a room for it to probably contain at least one pair of individuals born on the same day of the same month? Intuitively, it might seem that the room had better be large. After all, if two random individuals walk into the empty room, whatever birthday

[9] National Research Council Committee on Identifying the Needs of the Forensic Science Community, Strengthening Forensic Science in the United States: A Path Forward 144 (National Academy Press, Washington D.C., 2009).

[10] Saks and Koehler, supra n 6, at 212–13. Professor Koehler is quoted as the source of the example in Simon A. Cole, *"Implicit Testing": Can Casework Validate Forensic Techniques?*, 46 Jurimetrics J. 117, 123 n.33 (2006).

[11] For other applications or discussions of the Birthday Problem in forensic science, see Committee on DNA Technology in Forensic Science: An Update, National Research Council, The Evaluation of Forensic DNA Evidence (National Academy Press, Washington, D.C., 1996); Kaye, supra n 6; David H. Kaye, *Trawling DNA Databases for Partial Matches: What Is the FBI Afraid Of?*, 19 Cornell J. L. & Pub. Pol'y 145 (2009); Mark Page, Jane Taylor and Matt Blenkin, *Uniqueness in the Forensic Identification Sciences—Fact or Fiction?*, 206 Forensic Sci. Int'l 12, 14–18 (2011).

Downloaded from http://lpr.oxfordjournals.org/ by guest on January 30, 2014

	1	*2*	*3*	*4*	*5*	*6*	*7*	*8*	*9*	*n=10*
1		•	•	•	•	•	•	•	•	•
2			•	•	•	•	•	•	•	•
3				•	•	•	•	•	•	•
4					•	•	•	•	•	•
5						•	•	•	•	•
6							•	•	•	•
7								•	•	•
8									•	•
9										•
n=10										

FIG. 1. The $n(n-1)/2 = 45$ distinct pairs of possibly matching birthdays of $n = 10$ individuals.

Individual Number 1 has, the probability of duplicating it in Individual Number 2 is only 1/365. The correct answer, however, is that only 23 people need to enter the room. At that point, the probability of duplicating a birthday exceeds 0.5.[12]

This number is surprisingly small because the probability of 1/365 pertains to a match with respect to a specific birthday, but there are many other ways to have duplicate birthdays. As new individuals enter the room, their birthdays need not match the first occupant's birthday to obtain a match to someone else's birthday. As the number (n) of people in the room grows, the number of 'pairs' of birthdays that might match grows as the square of n. This geometric growth soon swamps the small probability per pair.

To see how this comes about, consider Fig. 1. It shows, for a mere $n = 10$ individuals, all n^2 pairs of birthdays that can be formed. Of course, the full square contains some equivalent pairs. For instance, (3,7) is the same as (7,3). In addition, the diagonal line of pairs such as (9,9) consists of every birthday compared to itself, which contributes nothing to the probability that two *different* individuals will have the same birthday. Thus, the total number of distinct pairs of birthdays of different individuals is half the number of the off-diagonal pairs: $(n^2 - n)/2 = n(n-1)/2$. For a large number n of birthdays, n and

[12] Consider the probability that the second person to enter the room will not have a birthday that matches the first person's birthday. This probability is 364/365. The probability that the third person will not match either of the previous two is 363/365. In general, the probability that n people have no birthday in common is

$$p = (364/365)(363/365)\cdots[(365-n+1)/365] = (1-1/365)(1-2/365)\cdots(1-(n-1)/365).$$

A Taylor expansion of $e^{-r/365}$ yields $1 - r/365 + (r/365)^2/2 - (r/365)^3/6 + \cdots$. When r is small compared to 365, the higher order terms are negligible, and $1 - r/365 \approx e^{-r/365}$. Using this approximation, we have

$$p \approx e^{-(1/365+2/365+\ldots+1-(n-1)/365)} = e^{-n(n-1)/(2\cdot 365)}.$$

The probability that some people in the room with n individuals have a birthday in common is then $1 - p \approx 1 - e^{-n(n-1)/(2\cdot 365)}$. When $n = 23$, $p = 0.500002$.

Downloaded from http://lpr.oxfordjournals.org/ by guest on January 30, 2014

$n - 1$ are practically the same. As such, there are approximately $n^2/2$ chances to duplicate a birthday. Because the total number of possibilities for duplication are not n, but $n^2/2$, the probability of some duplicated birthday even in a small group is much larger than one might have guessed.

In the Texas hypothetical, guns play the role of individuals, and the marks they leave on bullets are their birthdays. Apparently, there is a population of all pairs of the 100 000 guns that were in town during the 10-year period, and the firearms examiners have randomly sampled 100 examiners × 10 pairs/examiner-day × 365 days/year × 10 years = 3 650 000 pairs out of these $n = 100\,000 \times 99\,999/2$ pairs.[13] The space of possible pairs of 100 000 guns is so huge—almost $(10^5)^2/2 = 5\,000\,000\,000$—that even '1000 "collective years" of forensic science experience' (the 3 650 000 examinations) traverses only a small fraction of the space. Specifically, the probability of randomly stumbling on one of the 100 unusual pairs in any one comparison is $p = 100/5\,000\,000\,000 = 2 \times 10^{-8}$. The probability of 'not' doing so is $1 - p$. The probability of not encountering a matching pair n times in a row—of finding none of the 100 special pairings of matching guns in the sample of $n = 3.65$ million—is therefore $(1 - p)^n$, which is approximately $1 - np = 1 - 3.5 \times 10^6 \times 2 \times 10^{-8} = 93\%$. Thus, the example illustrates a serious weakness in the IAI-SWGFAST argument. As in the Birthday Problem, attending to all pairs of the elements of a set makes an enormous difference. Just as the probability of a duplicate birthday in a small set is more probable than one might have thought, the failure to falsify the postulated law of nature in a large set is far more probable than one might have surmised even though the relevant features are not unique.

3. Source attribution without uniqueness

The Texas Gun Case undermines the assumption of uniqueness in a large set, but it is not much of an argument against source attribution *per se*. Given the hypothetical data—zero observed matches out of 3.65 million observations—what can we legitimately infer from a finding that the markings on a bullet correspond to those on test bullets fired from a particular gun? How probable would it be to find such similar marks if the test bullets came from a different gun? Would it be misleading for an examiner to report that inasmuch as all the guns examined over the past 10 years have been unique, it is 'practically certain' that the bullet came from the gun that was tested?

The casual reader of the Texas Gun hypothetical might think that an examiner has 'little basis' to be confident in the source attribution. Yet, this is not quite what Professors Saks and Koehler wrote. Their assertion is that 3.65 million comparisons supply 'little basis for drawing conclusions about whether gun barrel markings, even in this single town, are unique'. While uniqueness implies that a match permits source attribution, a source attribution could be correct even if the features of the gun are duplicated within the population. In fact, we shall see that, on the facts of the hypothetical case, the source attribution is nearly certain to be correct. Thus, the problem with the theory of general uniqueness in the Texas town is not that it produces many false source attributions. Rather, it is that it leads the examiner to overstate the power of the source attribution, presenting it as a certainty when it is (on these facts) only a near certainty.

[13] Presumably, the examiners put the pair of guns that they tested back into the pool for future testing; otherwise they could have exhausted the town's supply of guns before doing millions of tests. We should also assume that the examiners keep track of which particular pairs they have sampled so that they do not do test the same pair twice. Sampling without replacement in this fashion introduces a dependence in the outcomes, but the effect is trivial. Finally, I assume that the examiners never misclassify an indistinguishable pair as distinguishable.

Downloaded from http://lpr.oxfordjournals.org/ by guest on January 30, 2014

To see this, we need to look beneath the Texas numbers. The hypothetical does not state that 100 guns out of the 100 000 in town always will be indistinguishable. It refers to 100 matching 'pairs'. At most, there could be as many as 100 non-overlapping pairs—rather like having a population of 100 000 individuals that included 100 identical twins. With identical triplets, quadruplets and quintuplets and so on, as well, there could be as few as 21 'duplicate' guns.[14]

Of course, even with only 21 problem guns out of 100 000, testimony that 'all' guns in the Texas town leave uniquely identifying marks on bullets, or that a particular gun has been 'individualized to the exclusion of all others', is always false. It invariably invites a jury to give the evidence more weight than it deserves. Suppose, however, that an expert testifies that the observed similarities in the striations on test bullets from a given gun and the bullet recovered from a crime scene establish a nearly certain association between the gun and the bullet. What is the probability that the reported association is false? The source attribution is false if the bullets came from the 'wrong' gun in one of the exceedingly rare pairs of matching guns.[15] In the worst scenario, there are 100 'twin' guns. The chance of stumbling on such a pair is only 100 out of 100 000, and in half of these cases, we will have hit upon the correct gun anyway. Duplication in the marks from guns is not unique in the Texas town, but the probability of a false association is, at worst, 50/100 000 = 1/2000. Thus, the Texas Gun Case does not demonstrate any large problem with an opinion that the bullet almost surely came from the gun that was examined.[16]

4. Alternatives to source attribution

The *Fingerprint Sourcebook* shows how wedded some forensic analysts are to the old paradigm of reporting an individualization based on a 'law of nature'. However, the expert community may be moving to more modest individualization testimony that acknowledges the possibility that indistinguishable marks from different objects exist.[17] Of course, before such limited source attributions are

[14] One 14-tuplet yields 91 distinct pairings, 1 quadruplet yields another 6, and 1 triplet gives the remaining 3 pairs. I am grateful to Jay Koehler for correspondence on this issue.

[15] Another route to a false association between a bullet and a gun is the failure to detect or properly assess the distinguishing features of the striations on a particular bullet. In that situation, the examiner could declare a match even when the bullets do not originate from a coincidentally matching gun. This possibility is a significant concern with source attributions, but it is not an objection to the uniqueness theory of individualization that the Texas Gun case is supposed to call into question. The hypothetical was offered solely to demonstrate that reliance on experience to establish uniqueness is highly fallible. The analysis here shows that this particular source of error does not substantially undermine the source attribution in the hypothetical case. Naturally, another source of error could have this effect.

[16] As a result, the Texas Gun Hypothetical does not seem to focus sharply on 'the practical implications for courts [rather than] more abstract and conceptual issues · · · with fewer implications for forensic science or judicial practice' (Koehler and Saks, supra n 6, at 1189) intended in Saks and Koehler, supra n 6.

[17] The chapter in SWGFAST's *Sourcebook* diverges from this government-sponsored group's recent elimination of the global 'exclusion of all others' language in its glossary (SWGFAST, Standard Terminology of Friction Ridge Examination, 23 March 2011, http://www.swgfast.org/documents/terminology/110323_Standard-Terminology_3.0.pdf). SWGFAST also plans to modify its *Standards for Conclusions*. ver. 1.0. 11 September 2003, available at www.swgfast.org/Standards_for_conclusions_ver_1.0.pdf. A draft guideline defines individualization as 'the decision by an examiner that there are sufficient features in agreement to conclude that two areas of friction ridge impressions originated from the same source'. Letter from Leonard G. Butt, Chairman, Scientific Working Group on Friction Ridge Analysis, Study and Technology, to Whom It May Concern (29 June 2010), available at http://www.swgfast.org/Comments-Positions/SWGFAST_NAS_Position_Clarification.doc. The classical approach to individualization for latent fingerprints also is rejected in NIST Expert Working Group on Human Factors in Latent Print Analysis, Latent Print Examination and Human Factors: Improving the Practice Through a Systems Approach, David H. Kaye ed. (National Institute of Standards and Technology, Gaithersberg, Maryland, 2012).

Downloaded from http://lpr.oxfordjournals.org/ by guest on January 30, 2014

allowed, controlled experiments should show that experts following the same procedure reliably can achieve high rates of accurate identifications.[18]

But to conclude that the Birthday Problem in the form of the Texas Gun case does not preclude source-attribution testimony is not to say that such testimony should be given. Unlike the creators of the hypothetical case, scientists do not know that there are exactly 100 pairs of duplicate guns in a town. Even so, firearms examiners have some data on the variability of marks left by the same gun and those left by different guns.[19] These experts could describe the nature and extent of the similarities and indicate, if only qualitatively, how much more likely it is to find such similarities when the same source as opposed to a different source produced the marks.[20] This 'likelihood ratio' or 'strength of evidence' approach to explaining the significance of forensic findings has a number of conceptual advantages over both source attribution and estimates of random-match probabilities.[21] It would leave it to the jurors to draw their own conclusion as to the hypothesis of a coincidental match along with other conceivable defence arguments and hypotheses, such as examiner error in measuring the similarities,[22] a plot to frame the defendant,[23] or some other reason to entertain reasonable doubt that the defendant committed the crime involving the gun.[24]

The Texas witness, not knowing the actual population proportion of duplicate pairs, also might offer an estimate of this figure. There are both Bayesian and frequentist methods for estimating the probability of an event that has occurred zero times in a sample of n observations—the 'zero-numerator' problem.[25] An approximate upper bound on the probability is $3/n$.[26] Here, the $3/n$ rule gives an estimate of not more than 3/(365 000 000)—about 8×10^{-7}, which is comfortably larger than the true value of 2×10^{-8}. In cases in which there have been fewer than 365 million comparisons, the

[18] Several experiments support this view. Bradford T. Ulery, R. Austin Hicklin, JoAnn Buscaglia, and Maria Antonia Roberts, *Accuracy and Reliability of Forensic Latent Fingerprint Decisions*, 108 PROC. NAT'L ACAD. SCI. 7733 (2011); Jason M. Tangen, Matthew B. Thompson, and Duncan J. McCarthy, *Identifying Fingerprint Expertise*, 22 PSYCH. SCI. 995 (2011); cf. B.T. Ulery R. Austin Hicklin, JoAnn Buscaglia, and Maria Antonia Roberts, *Repeatability and Reproducibility of Decisions by Latent Fingerprint Examiners*, 7 PLoS ONE (2012).

[19] See Ronald G. Nichols, *Defending the Scientific Foundations of the Firearms and Toolmark Discipline: Responding to Recent Challenges*, 52 J. FORENSIC SCI. 586 (2007) (describing studies).

[20] On this and other various modes of presentation, see Kaye et al., supra n 5; NIST EXPERT WORKING GROUP ON HUMAN FACTORS IN LATENT PRINT ANALYSIS, supra n 17. For a hypothetical transcript of such testimony, see David H. Kaye, *Likelihoodism, Bayesianism, and a Pair of Shoes*, 53 JURIMETRICS J. (forthcoming 2012).

[21] *Ibid.*

[22] Contra Jonathan J. Koehler, *Why DNA Likelihood Ratios Should Account for Error (Even When a National Research Council Report Says They Should Not)*, 37 JURIMETRICS J. 425 (1997); Jonathan J. Koehler et al., *The Random Match Probability in DNA Evidence: Irrelevant and Prejudicial?*, 35 JURIMETRICS J. 201 (1995) (contending that random-match probabilities are irrelevant and prejudicial when the risk of an erroneously ascertained match is much greater than the risk of a coincidental match).

[23] Laurence Tribe, *Trial by Mathematics: Precision and Ritual in the Legal Process*, 84 HARV. L. REV. 1329 (1971).

[24] Unfortunately, quantitative likelihood-ratio testimony, like quantitative random-match probabilities, is easily but fallaciously transposed into a 'source probability'. *McDaniel* v. *Brown*, 130 S.Ct. 665, 671 (2010) (*per curiam*). For examples, see Kaye et al., supra n 5, § 14.2.2.

[25] For example J.A. Hanley and A. Lipp-Hand, *If Nothing Goes Wrong, Is Everything All Right? Interpreting Zero Numerators*, 249 JAMA 1743 (1983); Thomas A. Louis, *Confidence Intervals for a Binomial Parameter after Observing No Successes*, AM. STATISTICIAN, 1981, at 154; Allan R. Sampson and Robert L. Smith, *Assessing Risks Through the Determination of Rare Event Probabilities*, 30 OPERATIONS RESEARCH 839 (1982); Robert L. Winkler, James E. Smith and Dennis G. Fryback, *The Role of Informative Priors in Zero-Numerator Problems: Being Conservative versus Being Candid*, AM. STATISTICIAN, 2002, at 1.

[26] B. D. Jovanovic and P. S. Levy, *A Look at the Rule of Three*, AM. STATISTICIAN, 1997, at 137. The 'counting method' with a confidence coefficient of 0.95 proposed for DNA typing in COMMITTEE ON DNA TECHNOLOGY IN FORENSIC SCIENCE, NATIONAL RESEARCH COUNCIL, DNA TECHNOLOGY IN FORENSIC SCIENCE (National Academy Press, Washington, D.C., 1992), is equivalent to this simple rule.

Downloaded from http://lpr.oxfordjournals.org/ by guest on January 30, 2014

estimates of the random-match probability will be larger. If a laboratory documented a zero numerator in 30 000 tests, it could infer a random-match probability of no more than 0.0001. For 3000 observations of distinct pairs, the upper bound on the random-match probability would be 0.001.[27]

Even though these probability estimates are generous in one respect,[28] there still is a risk that the jury will draw too strong a conclusion from such a probability. Perhaps this problem can be handled by cautioning the jury against inferring uniqueness or misconstruing the random-match probability as a source probability.[29] In any event, this concern is not an argument against the statistical soundness of the estimate. The use of sampling or probability theory here is elementary, but the quality of the estimate depends on the accuracy of the reported fraction of matching comparisons, the conditions under which the matches were made, and the analogy between a random sample and the casework or research sample.[30] Such is the nature of empirical reasoning.

5. Conclusion

The Texas Gun Case deploys the logic of the Birthday Paradox to show that even large samples cannot easily distinguish between a very rare event and a non-existent one. The claim endorsed by the IAI and SWGFAST that global, general uniqueness follows from a failure to find any duplicates in the experience of case analysts therefore is less impressive than it might seem to be at first blush. But the same mathematics indicates that when a random match probability is very small, the absence of uniqueness is not a barrier to all source-attribution testimony based on probabilistic reasoning. It simply means that individualization should not be based on the strong theory of global, general uniqueness.

Now that the unsatisfactory nature of the classical theory is becoming more widely appreciated, forensic science witnesses must use other explanations if they wish to continue providing source-attribution testimony instead of strength-of-evidence statements. The only plausible rationales recognize that source attributions—even those made by highly skilled experts in comparing patterns—cannot be guaranteed to be correct. Like any other non-trivial classification procedure, these categorical decisions entail some probability of false positives and negatives.

At the same time, just as advocates of global, general uniqueness theories cannot convincingly rely on the label 'individual' as opposed to 'class' characteristics to prove their claims, the advocates of a

[27] These computations do not account for misclassifications. See S. D. Walter and L. M. Irwig, *Estimation of Test Error Rates, Disease Prevalence and Relative Risk from Misclassified Data: A Review*, 41 J. Clinical Epidemiology 923 (1988).

[28] The 'Rule of Three' is a very conservative estimate. It corresponds to starting with the belief that p is equally likely to be anywhere in the interval between 0 and 1. P. Bickel and K. Doksum, Mathematical Statistics (1977). In light of general knowledge of the processes that produce toolmarks, fingerprints, etc., it is unrealistic to think that the probability that bullets fired from different guns rarely will show bigger differences than those fired from the same guns, that latent prints from the same fingers will have bigger differences than prints from different fingers, and so on. If so, the probability mass should be concentrated closer to 0 to begin with.

[29] On the problem of the transposition of the random-match probability, see *McDaniel* v. *Brown*, 130 S.Ct. 665 (2010) (*per curiam*); Kaye et al., supra n 5, § 14.1.2; David H. Kaye, *"False, But Highly Persuasive": How Wrong Were the Probability Estimates in* McDaniel v. Brown?, 108 Mich. L. Rev. First Impressions 1 (2009). A related concern is that the jury will overlook the probability of a misjudgement about the markings under examination. Cf. Jonathan J. Koehler, *Why DNA Likelihood Ratios Should Account for Error (Even When a National Research Council Report Says They Should Not)*, 37 Jurimetrics J. 425 (1997); Jonathan J. Koehler, *Error and Exaggeration in the Presentation of DNA Evidence at Trial*, 34 Jurimetrics J. 21 (1993).

[30] See generally David H. Kaye and David A. Freedman, *Reference Guide on Statistics, in* Reference Manual on Scientific Evidence 211–302 (Committee on the Development of the Third Edition of the Reference Manual on Scientific Evidence, Federal Judicial Center and National Research Council ed., 3rd edn. National Academy Press, Washington, D.C., 2011); Kaye, supra n 6.

Downloaded from http://lpr.oxfordjournals.org/ by guest on January 30, 2014

rule of law that would exclude all source-attribution testimony cannot convincingly claim that, by definition, only a census can establish uniqueness in any population.[31] They must rely instead on arguments of psychology and law regarding the mode of presentation that would best assist a judge or jury.[32] Sweeping assertions about the nature of science, epistemology and probability are less helpful here.[33]

Acknowledgements

A version of this article was presented at the Eighth International Conference on Inference and Forensic Statistics, Seattle, Washington, 21 July 2011.

Downloaded from http://lpr.oxfordjournals.org/ by guest on January 30, 2014

[31] Although Professors Koehler and Saks prefer to focus attention on 'individualization' involving the entire world, it is not clear whether they have abandoned the claim that as long as there is a non-zero probability of duplication in a population—no matter how closely it approaches zero—assertions of uniqueness in that population are scientifically impermissible. Compare Kaye, supra n 6 (offering an example in which only part of a population is sampled, exactly one member of the sampled part of the population has the characteristic in question, and it is obviously highly probable that no other member of the untested part of the population has the characteristic), with Koehler and Saks, supra n 6 (not responding to the example and conceding only that '[s]mall, closed population examples "work" only because one can compare target latent prints to every member of the potential source population ··· [and] all but one print can be eliminated as potential sources of the target latent.').

[32] For such arguments, see, for example, Kaye et al., supra n 5; Koehler and Saks, supra n 6.

[33] For assertions of this variety, see Vanderkolk, supra n 8; Saks and Koehler, supra n 6.