Painting the target around the matching profile: the Texas sharpshooter fallacy in forensic DNA interpretation†

WILLIAM C. THOMPSON*

Department of Criminology, Law and Society,
University of California, Irvine, CA 92697, USA

[Received on 17 November 2008; revised on 19 April 2009; accepted on 24 April 2009]

Forensic DNA analysts tend to underestimate the frequency of matching profiles (and overestimate likelihood ratios) by shifting the purported criteria for a ‘match’ or ‘inclusion’ after the profile of a suspect becomes known—a process analogous to the well-known Texas sharpshooter fallacy. Using examples from casework, informal and naturalistic experiments, and analysts’ own testimony, this article demonstrates how post hoc target shifting occurs and how it can distort the frequency and likelihood ratio statistics used to characterize DNA matches, making matches appear more probative than they actually are. It concludes by calling for broader adoption of more rigorous analytical procedures, such as sequential unmasking, that can reduce the sharpshooter fallacy by fixing the target before the shots are taken.

Keywords: DNA evidence; frequency; likelihood ratio; fallacy; sequential unmasking; DNA profile; bias; error; statistics; ACE-V.

1. Introduction

When evaluating the significance of scientific data, it is often helpful to calculate the probability that a particular event occurred by chance. Calculations of this type can be misleading, however, when they focus too narrowly on a given outcome without considering the broader context. For example, an epidemiologist who observes a ‘cancer cluster’ in a particular neighbourhood might compute the probability that random chance would produce so many cancer cases in that neighbourhood. This computation would be misleading because random processes are very likely to produce ‘cancer clusters’ in some neighbourhoods even though a cluster is unlikely in any particular neighbourhood (Neutra, 1990; Rothman, 1990; Thomas et al., 1985).

The Texas sharpshooter fallacy is the name epidemiologists have given to the tendency to assign unwarranted significance to random data by viewing it post hoc in an unduly narrow context (Gawande, 1999). The name is derived from the story of a legendary Texan who fired his rifle randomly into the side of a barn and then painted a target around each of the bullet holes. When the paint dried, he invited his neighbours to see what a great shot he was. The neighbours were impressed: they thought it was extremely improbable that the rifleman could have hit every target dead centre unless he was indeed an extraordinary marksman, and they therefore declared the man to be the greatest sharpshooter in the state. Of course, their reasoning was fallacious. Because the sharpshooter was

† Presented as part of the Seventh International Conference on Forensic Inference and Statistics, The University of Lausanne, Switzerland, 21st to 23rd August, 2008.

* Email: william.thompson@uci.edu
able to fix the targets after taking the shots, the evidence of his accuracy was far less probative than it appeared. The kind of post hoc target fixing illustrated by this story has also been called painting the target around the arrow.

In this article, I will argue that a process analogous to the Texas sharpshooter fallacy sometimes occurs in the production of forensic DNA evidence. Analysts create the impression that a DNA ‘match’ is a very small target that is unlikely to be hit by chance. But the probability of a coincidental ‘match’ may actually be much higher than analysts claim because a ‘match’ is not always a fixed target. Using examples from casework, informal and naturalistic experiments, and analysts’ own testimony, I will show that forensic DNA analysts sometimes shift their criteria for a ‘match’ based on the DNA profile of the suspect—in effect, they move the ‘target’ after the shots are fired. I will show how this post hoc target shifting occurs and how it can distort the frequency and likelihood ratio statistics used to characterize DNA matches, making matches appear more probative than they actually are. I will conclude by calling for broader adoption of more rigorous analytical procedures, such as sequential unmasking, that can reduce the sharpshooter fallacy by fixing the target before the shots are taken.

2. Hitting the target with DNA evidence: an easy case

The fallacy does not occur in every case. In many cases, DNA test results are clear and easy to interpret, which reduces opportunities for post hoc ‘target shifting’. Figure 1 shows STR test results for a crime scene sample (top electropherogram) and four suspects (lower four electropherograms) in what I will call an easy case. Each electropherogram shows the alleles found in the sample at three STR loci. The position of the ‘peaks’ on the electropherogram indicates which alleles are present. A computer program labels each peak with a number, which identifies the allele it represents. In order for two samples to ‘match’, they must have the same alleles at each locus.

Quick examination of these results shows clearly that Suspect 3 ‘matches’ the target profile and the other suspects do not. Hence, Suspect 3 is incriminated as a possible source of the bloodstain at the crime scene. If we think of the alleles in the evidentiary sample as the ‘target’, then it is apparent that Suspect 3 ‘hit’ the target exactly.

Figure 1. Electropherograms in an easy-to-interpret case.
To determine the probative value of this evidence for proving that Suspect 3 is the source of the bloodstain, we must consider the probability of hitting this target by chance. If a ‘match’ requires an exact one-to-one correspondence between the alleles in the two samples, as appears to be the case here, the computational procedure is straightforward. The first step is to determine the frequency (in some relevant reference population) of the matching alleles. The allele frequencies are then combined to determine the frequency of the genotype (pair of alleles) observed at each locus. If the genotype is heterozygous (i.e. the donor inherited a different allele from each parent, hence two different alleles are detected) then the genotype frequency is typically computed using the formula $2pq$, where $p$ and $q$ represent the frequencies of two heterozygous alleles. If the genotype is homozygous (i.e. the donor inherited the same allele from each parent, hence only one allele is detected), then the genotype frequency is computed using the formula $p^2$, where $p$ is the frequency of the homozygous allele (Butler, 2005, pp. 498–500). Once the genotype frequencies for the matching loci are computed, they are multiplied together to obtain the estimated frequency of the matching profile.

The frequency of the matching profile is then used to characterize the value of the DNA evidence for incriminating the matching suspect. It is often called the ‘random match probability’ and is described, to the trier-of-fact, as the probability that a randomly chosen unrelated person who was not the source of the evidentiary sample would happen to match. One might say that the profile frequency represents the probability of hitting the target by chance, where the target is an incriminating result.

The Texas sharpshooter fallacy does not appear to be a problem when computing the random match probability for Suspect 3 (Fig. 1) because the target was ‘fixed’. There was one and only one profile that would ‘match’ and the analyst could not change the requirements for a ‘match’ to fit a particular suspect.

3. Opportunities for target shifting: a harder case

Not all cases are as clear and easy to interpret as the case shown in Fig. 1. In cases I have reviewed over the past few years, evidentiary samples from crime scenes often produce incomplete or partial DNA profiles. Limited quantities of DNA, degradation of the sample, or the presence of inhibitors (contaminants) can make it impossible to determine the genotype at every locus (Butler, 2005, p. 168; Buckleton and Gill, 2005) In some instances, the test yields no information about the genotype at a particular locus; in some instances, one of the two alleles at a locus will ‘dropout’ (become undetectable). When testing samples with very low quantities of DNA, spurious alleles (i.e. alleles not associated with the sample) are sometimes detected, a phenomenon known as allelic ‘drop-in’

A further complication is that evidentiary samples are often mixtures of DNA from more than one person. It can be difficult to tell how many contributors there were to a mixed sample (Paoletti et al., 2005), and even more difficult to tell which alleles are associated with which contributor, hence at each locus there may be a number of different genotypes that a contributor could have, even if one assumes that all alleles have been detected. If, as is often the case, the analyst cannot be

---

1 To address concerns about population structure and inter-relatedness most laboratories introduce a slight adjustment to these simple formula that is known as a theta correction (Butler, 2005, pp. 506–507). For ease of exposition, I will ignore the theta correction as it makes little difference to the issues raised here.
certain that all alleles of all contributors have been detected, then the number of possible contributor genotypes is even greater.

To avoid false exclusions, analysts must make allowances for these phenomena when determining whether a suspect should be ‘included’ or ‘excluded’ as a potential contributor. As discussed below, these allowances not only expand the target, they create the potential for ‘target shifting’ and do so in ways that are not always taken into account by the frequencies (and random match probabilities) computed by most forensic laboratories. Consequently, the statistics presented in these cases often understate the actual probability of a coincidental ‘inclusion’. Figure 2 shows STR test results in what I regard as a harder case to interpret. This electropherogram shows the alleles detected in an evidentiary sample (assumed to be saliva) that was swabbed from the skin of a sexual assault victim. Below each allele are two boxes containing numbers. The upper number identifies the allele and the lower number indicates the peak height in relative florescent units. The height of the peaks corresponds to the quantity of DNA recovered from the sample.

The table at the bottom of Fig. 2 shows the profiles of four possible ‘defendants’. I will argue that it is not so clear which of these ‘defendants’ should be included or excluded as possible contributors.

The peak heights in Fig. 2 are much lower than those in Fig. 1 because a relatively small amount of DNA was recovered from the skin swab. For a sample of this type, the analyst must consider the likelihood of allelic dropout and the possibility of allelic drop-in. For example, at locus D3S1358 (hereafter D3), the analyst must determine whether the peak labelled 12 represents a true allele and, if so, whether it is associated with (i.e. from the same contributor as) allele 17; at locus FGA, the analyst must determine whether the peak labelled ‘OL Allele?’ is a true allele or an artefact. To compute the probability of a coincidental inclusion, the analyst must also consider whether this is a single source sample or a mixture.

2 ‘OL’ stands for ‘off-ladder’. In STR test kits, the alleles typically observed at each locus are called ‘ladder alleles’ because the test incorporates a control, called a ladder, which contains each of these common variants. An off-ladder allele may represent an unusual genetic variant. Alternatively, it could indicate a problem with the test, such as the presence of a spurious peak caused by a technical problem. Most experts who have looked at the electropherogram in Fig. 2, believe the ‘OL allele’ at locus FGA is an ‘artefact’—i.e. a spurious result, rather than a true allele. As noted below, however, experts differ about whether this spurious result might be ‘masking’ a true allele. And the masking theory tends to be invoked selectively depending on whether doing so is necessary to ‘include’ the ‘defendant.’
In cases I have reviewed in the USA, analysts almost always have had full knowledge of the profiles of the suspect (or suspects) when making determinations of this kind. The widespread failure to adopt ‘blind’ procedures for making these determinations sets the stage for target shifting. For example, if the suspect has the profile of Tom, the analyst might be more inclined to dismiss the 12 peak at locus D3 and the OL peak at locus FGA as artefacts than if the suspect has the profile of Dick.

4. An inadvertent experiment

As evidence for these assertions, let me describe an informal and rather inadvertent experiment that I have conducted over the past few years. This experiment was neither rigorous nor well controlled and the results were recorded in a manner that was neither blind, nor systematic, nor (I fear) entirely objective. Nevertheless, I think the experiment is worth describing here because it so nicely illustrates how something like the Texas sharpshooter fallacy can arise in the context of forensic DNA testing.

It began when I was invited to speak to a group of forensic DNA analysts at a meeting of the California Association of Criminalists. I was talking about the way forensic analysts interpret DNA evidence in ‘problematic cases’ and the tone of my comments was rather critical. I showed a slide in which the evidentiary profile in Fig. 2 was juxtaposed with a profile labelled ‘Defendant’. The ‘defendant’ profile was that of ‘Tom’ in Fig. 2.

I suggested that there might be some uncertainty about the inclusion of the ‘defendant’ because his profile did not account for all peaks in the evidentiary profile. In particular, neither the defendant nor the victim had the 12 allele at locus D3 nor the OL allele at locus FGA. How could we be sure, I asked, that the true contributor did not have genotype 12,17 at locus D3?

At that point, several analysts interrupted my talk to say that I did not know what I was talking about. One of the analysts stated, with the apparent support and affirmation of many others, that it would be obvious to any qualified expert that the 12 peak at locus D3 and the OL peak at FGA are artefacts that should simply be ignored when comparing the evidentiary profile to the suspect. The 12 peak could not represent a true allele, one of them said, because it ‘did not have the morphology of a true allele’. Even if it was a true allele, another said, it could not be from the same person who contributed the 17 allele because the ‘peak height disparity’ was too great for the peaks to have originated from the same individual. The OL Allele at locus FGA was ‘a known artefact’ that any competent analyst would recognize and ignore. Although I was a bit doubtful about some of these pronouncements, I felt that my arguments had been soundly defeated, particularly when a comment from the audience that characterized my position as ‘slimy’ and ‘irresponsible’ invoked enthusiastic applause.

Later that same month, I was scheduled to again speak about DNA evidence to an organization that called itself the Skeptics Society. As the date approached, the Skeptics informed me that a group of DNA analysts from a local crime laboratory and several local prosecutors who were experts on DNA evidence would be attending my talk and had asked for time on the program to respond to my comments. This seemed a bit intimidating but it also suggested an idea.

---

3 I did not set out to do an experiment, merely to create illustrations for teaching purposes. Over time I realized, however, that the reactions of the audience to variations in these presentations was rather telling.

4 The ‘morphology of a true allele’ entailed certain characteristics of symmetry and height–width ratio that were said to be recognizable to a trained analyst.

5 My approach to the talk was inspired in part by James Randi (‘The Amazing Randi’), who had preceded me as a speaker at the Skeptics Society. Randi is famous for devising techniques (often involving deception) to expose people who falsely claim to have supernatural abilities.
When I spoke, I again presented the evidentiary profile shown in Fig. 2. This time, however, the profile labelled ‘Defendant’ was that of Dick rather than Tom. I suggested that the inclusion of Dick was rather problematic due to uncertainty about whether the 12 peak at locus D3 was a true allele and because no 20 peak had been detected at locus FGA. I then invited the experts in the audience to tell me if I was wrong.

They wasted little time in doing so. One analyst told me that she could tell the 12 peak at D3 was a true allele based on its ‘morphology’. The ‘peak height disparity’ was nothing that would trouble a knowledgeable analyst, I was told, because discrepancies of that sort are expected with low quantities of DNA due to ‘stochastic effects’. The OL allele at locus FGA was indeed an artefact, but it could easily have ‘masked’ an underlying 20 allele. One analyst said the shape of the artefactual OL peak supported the theory of an underlying 20 allele. I was again condemned as ‘irresponsible’ for questioning the validity of this match. I then had the great pleasure of suggesting that the defendant’s profile might actually have been that of Tom rather than Dick, at which point my critics became noticeably less certain of the correctness of their interpretations (and the incorrectness of mine).

Encouraged by this felicitous result, I extended the experiment a few months later while speaking to an organization called the Association of Forensic DNA Analysts and Administrators. In that talk, the profile labelled ‘Defendant’ was that of Harry in Fig. 2. I suggested that the inclusion of the ‘defendant’ was problematic because the defendant’s genotype at D3 was 14,17, while the evidentiary profile had peaks at 12 and 17. I also noted the failure to detect the defendant’s 20 peak at locus FGA. Once again, the DNA analysts in the audience told me they saw no problem at all with the inclusion of the defendant (Harry). The failure to detect the defendant’s 14 allele at locus D3 could easily be due to allelic dropout and there might well be a 20 allele at locus FGA that was masked by artefact. The peak labelled 12 at locus D3 was an obvious artefact. If I were truly an expert like themselves, I was told, I would realize that my concerns about the inclusion of the defendant were groundless.

At this point, I wonder how much I would need to change the ‘defendant’ profile to get forensic DNA analysts to agree that the defendant should have been excluded. My friend Dan Krane from Wright State University presented this case to forensic analysts using the ‘defendant’ profile labelled Sally. Even with this profile, analysts have still insisted that the ‘defendant’ cannot be excluded. To reach this conclusion, they posited that the evidentiary sample consisted of a mixture of two contributors, one of whom has the 15 allele at locus vWA and the other who has the 17 allele. In other words, a new theory of the evidence (that it is a mixture) was invoked to ‘include’ Sally. Interestingly, the mixture theory was never mentioned when the ‘defendant’ had the other profiles.

5. The absence of formal standards for inclusion

These informal experiments can be criticized on a number of grounds, but the basic point they illustrate should be self-evident to anyone who looks closely at the actual practices of forensic laboratories. A key source of problems is the absence of any formal standards for distinguishing ‘inclusions’ from ‘exclusions’. I have looked carefully at the protocols of a number of laboratories in the USA and I am yet to see a protocol at any laboratory that sets forth standards that would allow an unambiguous determination of whether Tom, Dick, Harry or Sally should be included or excluded as contributors. Some protocols specify factors for analysts to consider in making such determinations, but none set forth objective standards and none require that any ‘guidelines’ that are mentioned be
followed consistently. Whether the comparison of profiles results in a finding of ‘inclusion’, ‘exclusion’ or ‘inconclusive’ is an entirely subjective determination.

In the absence of objective standards for distinguishing inclusions from exclusions, estimates of the probability of a coincidental inclusion are problematic. How can we estimate the percentage of the population who would be ‘included’ if the standards for inclusion are ill-defined and can be stretched one way or another by the laboratory? Estimating the size of the ‘included’ population under these circumstances is analogous to estimating the length of a rubber band. Just as a rubber band may be longer or shorter, depending on how far one is willing to stretch it, the size of the ‘included’ population may be larger or smaller, depending on how leniently or strictly the analyst defines the criteria for an ‘inclusion’. Because it is unclear just how far the laboratory might stretch to ‘include’ a suspect, the true size of the ‘included’ population cannot be determined. This is a fundamental problem that has not yet been resolved (and is not yet fully recognized) by the DNA testing community. This problem is particularly important in what I have called the hard cases in which there are incomplete and mixed profiles.

The standards for inclusion and exclusion are not just ill-defined and subjective, they are also flexible. As the Tom, Dick and Harry example suggests, the standards may shift to encompass the profile of suspects. This process may well occur without analysts being aware of it. Perhaps the process of comparing the defendant’s profile to the evidentiary profile channels analysts’ thinking about how the evidentiary profile might have arisen. Psychological studies have shown that the process of imagining how a particular outcome might occur increases peoples’ estimates of the probability that the imagined outcome will (or did) occur (Anderson et al., 1980). Hence, merely thinking about how the defendant’s DNA might have produced the observed profile could increase the analysts’ confidence that the defendant was a contributor. Moreover, while focusing on a defendant’s profile, analysts might also ignore, discount or simply fail to imagine other ways in which the same data might have arisen if the defendant was not a contributor.6

6 How the target shifts

To further illustrate the process of target shifting and its effects on statistical estimates, let us return to Tom, Dick, Harry and Sally. Consider first how an analyst might estimate the probability of an ‘inclusion’ at locus D3. When the defendant was Tom, the analysts tended to dismiss the 12 peak as an artefact, which led them to think that the perpetrator was a homozygous 17 (like Tom). Indeed, most of the analysts with whom I discussed the case of Tom told me they thought the correct formula for estimating the probability of a coincidental inclusion at D3 was \( p^2 \), where \( p \) is the frequency of the 17 allele. This would be the correct estimate if it were the case that the laboratory would only include someone with genotype 17,17 and would exclude anyone with a different genotype. But as the experiment shows, the ‘target’ is actually much wider than this. Neither Dick, Harry nor Sally had genotype 17,17, yet they were still ‘included’. So, an analyst who computed the match probability as \( p^2 \) would be underestimating the frequency of a coincidental inclusion; the analyst would, in effect, have painted the target around the arrow.

6 The notion that interpretation of scientific data can be influenced by extraneous factors, including the knowledge, assumptions and expectations of the scientist, is widely accepted (Risinger et al., 2002; Thompson, 2009). It is a major theme in science and technology studies (Hess, 1997) and the sociology of science (Longino, 1990) and has been extensively documented in psychological studies (Nisbett and Ross, 1980; Gilovich, 1991).
When Dick was the defendant, the analysts concluded that the perpetrator was likely a heterozygous 12,17 (like Dick). This led some of the analysts to say that the correct formula to use for estimating the probability of a coincidental inclusion was $2pq$, where $p$ and $q$ are the frequencies of the 12 and 17 alleles. This is a slightly wider target than $p^2$, but still not wide enough because this target excludes both Tom and Harry (who were ‘included’ when they were ‘the Defendant’).

Some laboratories are more cautious and, in a case like this, would estimate the probability of a coincidental match at locus D3 using the formula $2p$, where $p$ is the frequency of the 17 allele. Under this interpretation, all that is necessary to be ‘included’ is that the defendant possesses the 17 allele. This creates a more appropriate ‘target’ because it would include Tom, Dick, Harry and Sally. Nevertheless, the uncertainty about what is required for an ‘inclusion’ still leaves room to doubt whether the conservative $2p$ formula is conservative enough. Suppose, for example, that we posit a defendant (let us call her ‘Jane’) who has genotypes 12,14 at D3, 15,15 at vWA and 20,25 at FGA. Is it a certainty that Jane would be excluded? I believe most forensic DNA analysts would conclude that the evidentiary sample might be a mixture to which ‘Jane’ could be a secondary contributor and therefore would not exclude Jane (even though she lacks the 17 allele at locus D3). Hence, I believe that even the $2p$ estimate, which forensic analysts regard as extraordinarily conservative, still underestimates the probability of a coincidental ‘inclusion’.

The tendency to underestimate the probability of a coincidental match can occur at every locus of a case like this, and hence can affect the match probability assigned to the overall profile in a multiplicative manner. Let us next consider locus vWA. Faced with a suspect like Tom, Dick or Harry, most laboratories would treat the evidentiary profile as a single-source sample and would therefore use the formula $2pq$ to compute the probability of a coincidental match. In other words, they would assume that a suspect must have both the 15 allele and the 17 allele to be ‘included’. But this formula is underinclusive because it does not include Sally. When faced with a suspect like Sally, the theory that the evidentiary profile is a mixture is invoked and the target shifts accordingly. In a case where Sally is the defendant, I believe most laboratories would compute the probability of a coincidental inclusion at locus vWA using the formula $p^2+2pq+q^2$, which assumes that each of the contributors had either genotype 15,15 (like Sally) or 15,17 or 17,17. This formula is sometimes called the ‘random man not excluded’ formula and is also touted as extremely conservative. Once again, however, I believe this ‘conservative’ approach is not conservative enough in a case like this because it assumes that all alleles of all contributors have been detected. As the Harry case illustrates, analysts are not necessarily willing to make this assumption when faced with a suspect who has an allele that was not detected in the evidentiary sample. Suppose, for example, that Tom had had genotype 15,18 at locus vWA. Would he have been excluded? I think not. I think most analysts would conclude that the evidentiary sample might be a mixture to which Tom could be one of the two contributors. The fact that he has an allele at vWA that was not detected in the mixture would be attributed to allelic dropout.

7. More evidence of target shifting (and underinclusive statistics)

I realize that, to this point, my argument rests heavily on extrapolations from the informal Tom, Dick and Harry experiment. Readers might wonder whether interpretation of DNA test results is really as flexible as this ‘experiment’ suggests and whether laboratories actually underestimate the probability of a coincidental inclusion as much as the discussion above implies. So, let me turn to a
second line of evidence that supports my basic thesis—that ‘target shifting’ by DNA analysts leads them to underestimate the true likelihood of a coincidental inclusion.

The second line of evidence comes from cases in which forensic laboratories have computed a random match probability in the absence of a suspect or before all suspects are presented. In other words, these are cases in which the laboratory must paint the target before all the shots are fired. If I am correct that laboratories use targets that are too small and engage in target shifting, then they should occasionally be caught out in such cases. They should occasionally refuse to exclude a suspect who fails to fit the small target that they drew when computing the frequency of ‘included’ profiles (thereby proving that the target was too small). That is exactly what I have observed in several cases.

One such case was that of a prison inmate who sought post-conviction DNA testing in an effort to prove he was innocent of rape. A forensic laboratory tested a semen sample found on the victim’s underwear and a reference sample from the victim. Based on that evidence, and without having tested the inmate, the laboratory inferred the profile of the rapist and generated a statistical estimate of the random match probability. Figure 3 shows the electropherograms of the semen sample (labelled ‘Panties, sperm’). Below each electropherogram is a chart showing the laboratory’s conclusions regarding the alleles in the panties/sperm sample and the profile of the rapist. The chart also lists the profile of the victim, which was determined in a separate assay. The bottom row in the chart shows the formula the laboratory used to estimate the random match probability in this case.

As discussed earlier, the formula used to compute the frequency defines the size of the target. By using the formula $2pq$ to compute the frequency of matching profiles at locus THO1, e.g. the laboratory is implicitly saying that only persons with a particular heterozygous genotype (7,9) will be included as potential sources of the male component of the sperm sample and, by implication, that persons with any other genotype will be excluded. And this assertion is the basis for the random match probability that will be used to characterize the value of the evidence if a match is found. If, as I believe, the laboratory is defining the target too narrowly, then it should be possible to produce a suspect who does not have the obligatory genotype but whom the laboratory nevertheless refuses to exclude.

That is what happened in this case. The inmate was tested after the laboratory inferred the profile of the rapist and computed the random match probability. His profile matched the inferred profile of the rapist at every locus except THO1. At that locus his genotype was 7,7. So what did the laboratory do? As in the case of Sally discussed above, the laboratory invoked the theory of an additional unknown contributor. Under this theory, the inmate could have contributed the 7 allele at THO1 while the 9 allele came from a second unknown contributor. Based on this theory, the laboratory decided that the inmate could not be excluded and he therefore remains in prison.

The failure to exclude the inmate makes it clear that the laboratory’s statistical estimates were underinclusive—i.e. that they understated the actual probability of a coincidental inclusion. As with the Tom, Dick and Harry case, however, it is unclear what the actual random match probability is.

---

7 The victim, whose genotype is 8,10, could not account for the 9 allele, although her DNA could explain the 10 allele if we also invoke the theory of allelic dropout to explain the failure to detect her 8 allele.

8 Interestingly, there is no indication in the laboratory notes or the laboratory reports that anyone suspected the sperm sample contained DNA of three people (the victim and two others) until after the inmate was tested. It seems rather unlikely that a random individual would have a profile that overlapped sufficiently with the victim and inmate that its presence was betrayed by only a single allele. To avoid excluding the inmate, then, the laboratory invoked a theory that, a priori, seemed rather unlikely. Like the target-shifting problem that I am discussing here, this practice also reduces the value of a DNA match in ways that are not reflected in standard random match probabilities, but that is a topic for another paper.
Fig. 3. Electropherograms of the evidentiary sample in a post-conviction case (with the laboratory’s inferences regarding the rapist’s profile and formulae for estimating the random match probability).
After testing the inmate, the laboratory revised its estimate of the random match probability by changing its frequency formula, but only for locus THO1. The laboratory determined that at THO1, a contributor could have any of the following genotypes: 7,9, or 7,?, or 9,?, where the question mark indicates any possible allele. (In other words, the laboratory assumed that the test may have failed to detect all of the contributors’ alleles at this locus.) Accordingly, the laboratory used the formula $2pq + 2p + 2q$ to compute the frequency of ‘included’ genotypes at THO1. At every other locus, however, the laboratory used the same formulae it had used before.

This second statistical estimate is obviously yet another example of painting the target around the arrow. If allelic dropout is possible at locus THO1, then it is also possible at other loci. Why then did the laboratory not take the possibility of allelic dropout into account when computing the frequency of matching genotypes at these other loci? The answer, of course, is that the laboratory only paints targets where it sees arrows. Would the laboratory really have excluded the inmate if he had, say, genotype 23,28 at locus FGA or genotype 31,31 at locus D21? I seriously doubt it. Consequently, I think the revised random match probability computed by the laboratory was still underinclusive.

How should the laboratory have computed the random match probability? That question is difficult to answer because, as discussed above, no one knows what the actual criterion for inclusion might be in a case like this. Unless laboratories can delineate in advance (i.e. without knowing the profile of suspects) what genotypes are ‘included’ and ‘excluded’ at each locus, and stick by those determinations, their statistical estimates are likely to be underinclusive in many if not most cases. The degree of the bias introduced by target shifting is difficult to estimate and undoubtedly varies from case to case. But the bias clearly exists and will continue to do so until laboratories adopt more rigorous procedures for interpreting test results.

8. DNA and indeterminacy

In the absence of clear standards for ‘inclusion’ and ‘exclusion’, different experts evaluating the same evidence may reach different conclusions: one may conclude that a particular suspect is ‘included’, while another concludes that the same suspect is ‘excluded’. An interesting example of such a disagreement arose in the case of Commonwealth v. Leon Winston, which was tried in the state of Virginia in 2003. Winston was charged with a horrendous double homicide in which a man and a woman were shot and killed in the presence of the woman’s young children during a drug-related robbery. Winston was convicted and sentenced to death.

One piece of evidence used to link Winston to the crime was his ‘inclusion’ as a possible contributor to DNA on a glove that was found near the murder scene. There was uncontested testimony that a man named Hardy was the owner of the glove and that he had loaned the glove and its mate to a man named Brown, who allegedly assisted Winston in the crime. The prosecution theory was that Brown gave the gloves to Winston, who wore them while committing the crime, and discarded them thereafter. A DNA analyst from the state crime laboratory testified for the prosecution that the left glove of the pair contained a mixed DNA profile in which Hardy, Brown and Winston were all included as possible contributors. However, an expert called by the defense, who reviewed the work of the state laboratory, reached a different conclusion on the same evidence. She testified that while Hardy and Brown should be included as possible contributors, Winston should be excluded.
Figure 4 shows the DNA profiles reported by the state laboratory. The laboratory used gel electrophoresis with a commercial test kit called Promega PowerPlex that examines 15 STR loci. However, the test produced results at only 10 of the 15 loci. Close examination of the alleles detected at those 10 loci, which are listed in Fig. 4, reveals the source of the disagreement between the experts. Although the mixed profile found on the glove contains a number of alleles consistent with the three putative contributors, some of their alleles are missing. The bold numbers in Fig. 4 identify alleles possessed by contributors that were not detected in the sample from the glove.

The missing alleles did not trouble the prosecution expert, who attributed their absence to ‘allelic dropout’. The defense expert agreed that allelic dropout could account for the absence of some of Brown’s and Hardy’s alleles, but apparently concluded that too many of Winston’s alleles were missing for him to be a plausible contributor and therefore decided that Winston should be ‘excluded’.

At this point in the history of DNA testing, it is impossible to determine which expert is right and which expert is wrong in disagreements of this type because, as noted above, there are no standards for distinguishing ‘inclusions’ and ‘exclusions’. Leading thinkers in the field have recognized this problem and have commented on how an expert might approach the analysis of evidence in cases where ‘allelic dropout’ might have occurred (Buckleton and Triggs, 2005; Gill et al., 2006). But these commentaries and suggestions have, to my knowledge, not yet been translated into rules of interpretation that an analyst can easily apply when interpreting evidence like that in the Winston case. In most laboratories in the USA, such determinations have been and remain a matter of subjective judgment.

9. Misuse of likelihood ratios

The experts in the Winston case also differed dramatically over the proper statistical characterization of the evidence linking the defendant to the glove. The state’s expert presented a likelihood ratio—declaring that:

The DNA profile I obtained from the sample from the glove is 1.8 billion times more likely if it originated from Leon Winston, Kevin Brown and David Hardy than if it originated from three unknown individuals in a Caucasian population. It’s 1.1 billion times more likely it originated from these three individuals than if it originated from three unknown individual [in] the black population; 2.9 billion times more likely to have originated from these same three individuals than if it originated from three unknown individuals in the Hispanic population. (Reporter’s Transcript, Vol X., p. 335–36).

9 The analyst chose to list some alleles in parentheses to indicate that the bands indicating those alleles are ‘less intense than types not in parentheses’. Figure 4 replicates the analyst’s reporting of those alleles.
Assuming the testimony was correctly transcribed, the expert appears to have botched the basic description of the likelihood ratios that she computed for the Black and Hispanic populations. The likelihood ratios she was attempting to explain describe how much more likely the evidence is under one hypothesis (Hardy, Brown and Winston were the contributors) than another hypothesis (three unknown persons were the contributors), not the likelihood that the evidence ‘originated from these three individuals’. After her reference to the Caucasian population, she seems to have inverted the relevant conditional probabilities.

The prosecutor later made matters worse by mischaracterizing (further) what the expert had said. He attempted to explain the expert’s statistical statement about the glove evidence as follows:

[S]he said that the odds of finding a random match unrelated to the defendant, just finding someone randomly that had those same characteristics and those same loci, is one in 1.1 billion in the black race. (Reporter’s Transcript, Vol XI, p. 176)

The prosecutor, it appears, converted the likelihood ratio (for the ‘black race’) into a random match probability.

Beyond these unfortunate mischaracterizations, there are several problems with the likelihood ratios that the state’s expert presented. First, the two hypotheses that the expert compared were neither relevant nor appropriate for the case. As noted above, it was uncontested that Hardy and Brown had had contact with the gloves. The real issue was whether Winston, rather than some unknown person, was the third contributor. Hence, the expert should not have asked how much more likely the observed evidence would be if the DNA on the gloves originated from Hardy, Brown and Winston than if it originated from three unknown persons; she should have asked how much more likely the evidence would be if it originated from Hardy, Brown and Winston than if it originated from Hardy, Brown and a random unknown person. By my calculations, the answer the expert would have reached to this latter more appropriate question (using data for African Americans) was 789, rather than 1 100 000 000. This is a big difference, although, as discussed below, I believe even the more modest likelihood ratio of 789 greatly overstates the value of this evidence for incriminating Winston.

10. Another example of target shifting

A second problem with the likelihood ratios computed by the state’s expert in the Winston case is that they failed to take full account of the evidence. When testing the glove, the expert obtained reportable results at the 10 loci shown in Fig. 4. However, the expert used data from only five of these loci when computing the likelihood ratios. The five loci she chose (FGA, D8, vWA, D18 and D5) were those for which all alleles of the three putative contributors had been detected. She elected to ignore the other five loci (Penta E, D21, THO1 and D3) where the test had failed to detect one or more alleles of the putative contributors.

Why did she choose to focus on only 5 of the 10 loci when computing statistics? The defense lawyer asked her about this during cross-examination. Her answers made it clear that there was nothing about the test results on glove sample that dictated this choice. She focused on those loci because those were the ones where she had found a match with the putative contributors. After the defense lawyer pointed out that neither of Winston’s alleles at locus D7 had been detected in the glove sample, the following exchange occurred:
Q: You didn’t include [locus D7] in your calculations, did you? A: No, because like I said before, in order to calculate statistics, I only chose the loci that had the three individuals I was including in—their complete profiles in. (Reporter’s Transcript, Vol X, p. 356)

In other words, this analyst only painted targets where she saw arrows. By dredging through her findings, the analyst risked ignoring data that tended to disconfirm the government’s theory of the case while focusing only on that portion of the data that tended to confirm the government’s theory. This is a cardinal sin of statistical inference.

Interestingly, the defense expert made a similar error. Rather than computing likelihood ratios, the defense expert computed ‘random man not excluded’ statistics. She testified that ‘within the African-American population, the probability of randomly selecting an individual who would be included in [the mixture found on the glove] at the five locations that were used in [the prosecution expert’s] analysis is 1 in 195’. (RT, p. 25) The frequency of 1 in 195 is obviously more modest than even the likelihood ratio of 789 mentioned above. But is it conservative enough? I think not because, once again, the analyst painted the target around the arrow.

When computing the random match probability, the defense expert, like the prosecution expert, focused only on the five loci where all alleles of the putative contributors had been detected. For those loci, she computed the random match probability by taking the sum of frequencies of all ‘included’ genotypes. For example, at locus vWA, where alleles 16, 17 and 20 were observed on the glove, the expert computed the sum of the frequencies of genotypes 16,16; 17,17; 20,20; 16,17 and 16,20, on the assumption that anyone with one of those genotypes would be ‘included’ as a potential contributor to the glove at that locus (and presumably that anyone with a different genotype would be excluded).

For the other five loci (where some alleles of the putative contributors were not detected), the defense expert set the random match probability to 1. Her rationale was that no one can be excluded at these loci due to the possibility of allelic dropout (undetected alleles). If no one can be excluded, then everyone is ‘included’, making the probability of a random match 1 (1.0) for that locus. Her logic seems sound as far as it goes. At locus D7, e.g. neither Winston nor Hardy was excluded even though they have alleles that were not detected in the mixture. If they could not be excluded, then it would appear that no one else could be either. No matter what genotype a suspect has, the suspect cannot be excluded if the analyst assumes the suspect’s alleles were not detected.

The problem, of course, is that the analyst’s conclusion about whether all alleles were detected at a locus appears to have been influenced by knowledge of the profiles of the putative contributors. If the alleles of Hardy, Brown and Winston were detected at a locus, then she assumed that all alleles of all potential contributors had been detected and computed the random match probability accordingly; if one or more alleles of a putative contributor were not detected, she assumed that there was allelic dropout and that no one could be excluded. This appears, then, to be another example of painting the target around the arrow.

As with the earlier examples, this procedure may have led the analyst to underestimate the frequency of people in the population who would have been included as potential contributors. Suppose, for example, that Winston’s genotype at locus vWA were 16,18 rather than 16,16. Would he have been excluded? I seriously doubt it. I believe the analyst would simply have assumed that the failure to detect the 18 allele was another instance of allelic dropout. In other words, I believe the occurrence of allelic dropout was inferred from the very data that it was invoked to explain. Hence, the
actual probability that a randomly chosen person would be included as a potential contributor was higher, perhaps much higher, than the defense expert’s estimate of 1 in 195.

I believe that the only real requirement for including Winston as a contributor was that he possessed two alleles observed in the mixture that could not have come from Brown or Hardy (allele 22 at locus FGA and allele 15 at locus D18). But approximately half the human population possesses those two alleles, so the value of this ‘match’ for incriminating Winston seems rather slight. It is indeed unfortunate, in a case where the defendant’s life was at stake, that the prosecutor told the jury the probability of a random match to an important sample was 1 in 1.1 billion. In my opinion, the actual random match probability is close to 1 in 2; hence, the number the prosecutor gave the jury may have understated the true value by approximately nine orders of magnitude.

11. Can likelihood ratios solve the problem?

In 2006, a DNA Commission of the International Society of Forensic Genetics offered a helpful analysis of mixture interpretation that included extensive discussion of the problem of allelic dropout (Gill et al., 2006). The Commission suggested that a likelihood ratio be computed for each locus. It suggested that the possibility of allelic dropout be addressed by incorporating a dropout probability, designated Pr(D), into the likelihood ratio. For a locus where the suspect has genotype ab, but only allele a is observed in the evidentiary sample, the Commission recommended a formula initially proposed by Buckleton and Triggs (2005):

\[
LR = \frac{Pr(D)}{p_a(p_a + 2Pr(D)(1 - p_a))},
\]

where \(p_a\) is the frequency of allele a.

Although this formula makes sense conceptually, it is difficult to apply even in simple cases due to uncertainty about the dropout probability, \(Pr(D)\). The Commission noted that many laboratories have carried out experiments that provide information relevant to \(Pr(D)\). In general, the probability of dropout of one allele in a genotype increases as the peak height of the other allele decreases—in other words, allelic dropout is more common when peak heights are low due to limited quantities of DNA. Empirical study may allow determination of a threshold peak height above which \(Pr(D)\) \(\approx\) 0, and this could provide a basis for exclusions when the evidentiary sample includes one of the suspect’s alleles above the threshold but not the other. However, estimation of \(Pr(D)\) when the observed peaks are below this threshold will inevitably require a considerable element of guesswork, as the relationship between peak height and \(Pr(D)\) is unlikely to be linear and \(Pr(D)\) may also vary based on factors other than peak height, such as the nature and condition of the samples or the substrates on which they are found.

Matters quickly become even more complicated when formulae incorporating \(Pr(D)\) are applied to mixtures. Overlapping and masking alleles may make it difficult if not impossible to determine whether one allele from a suspect is above or below the level where the other is likely to dropout. Consider, e.g. locus PentaE in the Winston case. Although a 12 allele consistent with Winston was found on the glove, the evidence for the presence of this allele could be accounted for, in part if not in whole, by the 12 allele of Hardy. What then is the probability that Winston’s 15 allele would dropout, if he is indeed a contributor? I suspect that any answer an expert might provide to this question based on a subjective evaluation would be no better than a wild guess, and that different experts might (with persuasive sounding justifications) provide a variety of different estimates. Hence, computation of \(Pr(D)\) will be problematic in such cases. Even under the best of circumstances, when
experts have adequate empirical data about dropout probabilities, the computations would need to take account of peak heights at multiple loci, whether the evidentiary sample is degraded, whether different components of the evidentiary sample might be degraded to different degrees and other variables of this nature. The Commission acknowledged that ‘[e]xpansion of these concepts to mixtures is complex and that is why they are not generally used’. My concerns about it go beyond mere complexity. I wonder whether these determinations fall in a realm of indeterminacy, where so many unmeasured variables can affect the results that dropout probabilities simply cannot be determined with any degree of accuracy.

12. Automated systems as a possible solution

One possible solution to the problem of target shifting is the use of automated (programmed) systems for profile interpretation (Gill et al., 2007; Perlin, 2003, 2006). These systems can deal with the complexity of the computations needed to incorporate dropout probabilities and other such variables into likelihood ratios. They also have the tremendous advantage of objectivity. Unlike human analysts, their interpretations of evidentiary profiles are not influenced by the profile of suspects, hence they resolve the problem of target shifting. Moreover, they may extract information from underlying genetic data more efficiently than human analysts. According to promoters, these systems generally produce more extreme (and impressive) likelihood ratios when evaluating forensic comparisons because they make more complete and efficient use of the available data than the simpler computational methods typically used by laboratories.

If these claims are true, then automated systems may well solve the problems raised in this paper. Of course, commercial claims of this type need to be carefully evaluated. Before accepting an automated system for use in criminal cases and before accepting evidence produced by such a system in court, we should demand to see a careful program of validation that demonstrates the system can accurately classify mixtures of known samples under conditions comparable to those that arise in actual forensic cases. The fact that an automated system can produce answers to the questions one puts to it is no assurance that the answers are correct. While automated systems appear promising, their ability to handle ‘hard cases’ like those discussed in this paper remains to be fully evaluated.

It is also important that the operation of these systems be sufficiently transparent to allow experts to evaluate their conclusions. There must be a way to look inside the ‘black box’ to determine whether its conclusions in a particular case are well grounded. Because these systems are proprietary, their underlying code is not readily available for examination. Nevertheless, there should be ways to interrogate the system about the premises and assumptions underlying the conclusions in a particular case. To generate a likelihood ratio in the Winston case, e.g. an automated system would need to assess such issues as the probable number of contributors, the relative amount of DNA contributed by each, the dropout probability at each locus and the degree of degradation. These intermediate assessments should be accessible so that other experts can evaluate whether they are reasonable. A system that works well in general may trip up and reach implausible intermediate conclusions in some cases. If there is no way to evaluate whether that happened, it will raise concerns about the suitability of the system for the legal arena.

13. Sequential unmasking as a possible solution

Another way to deal with the problem of target shifting is simply to use more rigorous procedures for interpretation—procedures that ‘mask’ or ‘blind’ analysts to the profiles of suspects when they
make critical judgments about the profiles of evidentiary samples. Krane et al. (2008) recently proposed a procedure for interpretation of DNA profiles that they called ‘sequential unmasking’. While the Krane et al. proposal was designed primarily to minimize observer effects in the interpretation of DNA profiles, sequential unmasking would have the additional benefit of preventing target shifting.

The sequential unmasking procedure is simple and could easily be implemented in any forensic laboratory. When interpreting a DNA test, analysts would begin by looking just at evidentiary samples. In the initial interpretation, the analyst would determine the alleles associated with each sample, assess the number of contributors and assess the likelihood at each locus that the test procedure failed to detect some of the alleles of contributors (i.e. the dropout probability). At this point, and before looking at any reference samples, the analyst would determine (and make record of) the genotypes that would cause a person to be included or excluded at each locus.

After this initial interpretation, the analyst would unmask information about the reference samples in a sequential manner, beginning with samples from expected contributors, such as the victim in a sexual assault case. After considering the profiles of expected contributors, the analyst would re-evaluate the evidentiary profile to determine the possible genotypes of unknown contributors. At this stage, and before looking at the profiles of any suspects, the analyst would compute the frequency (in appropriate reference populations) of individuals who would be included as possible contributors. By so doing, the analyst would be documenting an unbiased assessment the genotypes of possible contributors. The analyst would, in effect, be fixing the target and computing its size. Only when these computations were recorded would the analyst undertake the final step of determining whether the suspects have the genotypes necessary to hit the target.

By forcing analysts to paint their targets before knowing where the shots hit the barn, sequential unmasking would prevent target shifting and lead to more accurate and appropriate estimates of random match probabilities. It is a simple and scientifically rigorous procedure that could and should be adopted by every laboratory.

Unfortunately, most laboratories in the USA have not yet adopted such procedures. Their failure to do so can be seen in laboratory protocols, in laboratory notes and in the testimony of analysts—who often make it clear, as in the Winston case, that their interpretation of the evidentiary sample depended, in part, on the profile of suspects. In my view, this is an endemic problem that affects a significant percentage of all forensic DNA cases. The problem may well have become worse in recent years as DNA testing has been used increasingly to analyze marginal samples that produce incomplete, ambiguous profiles like those in the Tom, Dick and Harry case. As this paper shows, target shifting in such cases can cause analysts to significantly underestimatereestimate random match probabilities and thereby overstate the value of DNA evidence. Forensic DNA laboratories can and should take more care to avoid this problem.

14. Beyond DNA

Now that we have seen how the Texas sharpshooter fallacy can affect the interpretation and characterization of forensic DNA tests, it is worth considering whether this fallacy might also influence other forensic disciplines, such as latent print and toolmark examination. While a complete discussion of this question is beyond the scope of the present article, I will offer a few observations.

First, DNA testing is not the only discipline in which analysts must rely on subjective judgment when determining whether two samples could have a common source. As a recent report of the
National Research Council (2009) has noted, ‘match’ determinations in the pattern comparison disciplines generally require ‘a subjective decision based on unarticulated standards’. In fact, there may well be more latitude for subjectivity in pattern matching disciplines than in DNA testing because DNA analysts look for the same set of features (alleles) in every sample, while the set of features used to make pattern matches, such as latent print and toolmark identifications, may differ in every case.

Second, in the pattern comparison disciplines, as in DNA testing, analysts typically fail to employ blind procedures. They know the characteristics of the reference samples and often are privy to information suggesting which reference samples are likely (or expected) to match particular evidentiary samples (Risinger et al., 2002). As noted earlier, the combination of subjective judgment with the failure to use blind procedures sets the stage for target shifting. In these circumstances, there is a danger that knowledge of the reference samples will influence analysts’ interpretations of the evidentiary samples and their decisions about which features are relevant and irrelevant to the comparison. For example, an examiner may decide to disregard features of an evidentiary sample that fail to match a reference sample when the analyst would credit and use those same features if they did happen to match.

A possible example of this process arose in the notorious Mayfield case, in which Federal Bureau of Investigation (FBI) examiners mistakenly identified a latent print associated with a terrorist incident in Madrid as having been made by Brandon Mayfield, a lawyer from Portland, Oregon (United States Department of Justice, Office of Inspector General, 2006; Stacey, 2005; Thompson and Cole, 2005). It was later determined that the source of the print was an Algerian suspect named Ouhnane Daoud (Thompson and Cole, 2005). There were discrepancies between the latent print and Mayfield’s print which the FBI examiners decided to ignore, apparently because they attributed them to distortion of the latent or the overlay of another print. Yet, these same features were deemed relevant and informative when they matched Daoud’s print. This example suggests that latent print examiners can engage in target shifting (perhaps without even realizing it), just as DNA analysts can. And the false match with Mayfield illustrates how target shifting can expand the match criteria in ways that create a danger of false matches—ways that analysts may not take fully into account when judging the probability of a coincidental match.

In some disciplines, most notably latent print examination, analysts employ the analysis-comparison-evaluation-verification (ACE-V) method, which requires identification of relevant features in the evidentiary print and an assessment of whether there is sufficient detail to allow a meaningful comparison (analysis) before the analyst examines the reference print to see if it contains similar features (comparison) (Ashbaugh, 1999, 1991; Haber and Haber, 2008). The ACE-V method may well prevent analysts from being influenced by a reference print during the initial evaluation of the evidentiary print (analysis phase), and hence may accomplish part of what the sequential masking procedure is designed to accomplish in DNA interpretation. However, ACE-V does little to prevent observer effects and target shifting once the analyst moves beyond the analysis stage, and therefore cannot be regarded as comparable to (nor an acceptable substitute for) a more rigorously blinded procedure like sequential unmasking.

The key innovation of sequential unmasking is a requirement placed on the analyst after analysing the evidentiary sample and before looking at the reference sample. At this point, the analyst must specify and record which features of a reference sample will lead to a determination of ‘match’.

---

10 The US Department of Justice Report (2006) suggested that the FBI examiners had engaged in ‘circular reasoning’, which is another way of saying that their interpretations of the evidentiary print were influenced by their knowledge of Mayfield’s reference print.
PAINTING THE TARGET AROUND THE MATCHING PROFILE

(inclusion) and which will lead to a determination of non-match (exclusion). This process ‘fixes the target’, thereby precluding the potential for target shifting and the sharpshooter fallacy. There is no comparable specification and recording of matching criteria in ACE-V. To the extent ACE-V ‘fixes the target’, the fix exists solely in the mind of the analyst. In light of the extensive evidence that observer effects can operate without conscious awareness (Risinger et al., 2002; Thompson, 2009; Krane et al., 2009), the effectiveness of a purely mental fix seems doubtful. As the National Research Council (2009, pp. 5–12) report notes, ‘ACE-V does not guard against bias; it is too broad to ensure repeatability and transparency; and does not guarantee that two analyst following it will obtain the same results’.

In sum, target shifting is likely to be a problem in the pattern-matching disciplines—a problem that can lead analysts to underestimate the likelihood of coincidental matches. ACE-V is not an adequate solution to this problem. As in the area of DNA analysis, more rigorous interpretive procedures, such as sequential unmasking, will be needed to prevent examiners from painting targets around their forensic arrows.

Funding

National Science Foundation (SES-0617672).

REFERENCES


