

Challenging Firearms and Toolmark Identification — Part Two

Firearms and Toolmark Identifications Cannot Be Made to a Reasonable Degree of Scientific Certainty

report by the National Research Council Committee to Assess the Feasibility, Accuracy, and Technical Capability of a National Ballistics Database (the NRC Report) explicitly recognized that firm statistical foundations do not exist for firearms and toolmark examination. "Conclusions drawn in firearms identification should not be made to imply the presence of a firm statistical basis when none has been demonstrated." According to the committee, the tendency of exam-

Editor's Note: Part One of this article was published in October 2008. In that installment, the author began her argument, continued here, that firearms and toolmark examiners lack a scientific basis for the claim that they can single out the particular gun that fired a particular ammunition component.

iners "to cast their assessments in bold absolutes, commonly asserting that a match can be made 'to the exclusion of all other firearms in the world' ... cloak[s] an inherently subjective assessment of a match with an extreme probability statement that has no firm grounding and unrealistically implies an error rate of zero."²

Although its criticism of "extreme probability statements" is welcome, the committee failed to appreciate the full, logical implications of its finding. If no "firm statistical basis ... has been demonstrated" for firearms and toolmark identification, there is no scientific basis for *any* testimony about firearms and toolmark matches. The problems do not disappear if, instead of testifying to a match to the "exclusion of all firearms in the world," examiners moderate their language.

In *United States v. Monteiro*, Judge Saris came closer to understanding the logical implications of the absence of statistical empirical foundations. She cited the explanation, in my Columbia Science and Technology Law Review article, of "the need for [an] adequate statistical empirical foundation to determine 'the likelihood that the toolmarks made by a randomly selected tool of the same type would do as good a job as the toolmarks made by the suspect tool at matching the characteristics of the evidence toolmark[.]"3 She then stated, "As of the writing of this opinion ... such a standard is not prevailing in the field, and an expert may not assert any degree of statistical certainty, 100 percent or otherwise, as to a match." Judge Saris failed, however, to draw the correct inference that the absence of any degree of statistical certainty makes firearms and toolmark identification testimony per se unreliable and inadmissible. Instead, she held that although a firearms and toolmark expert may not testify that there is a match to an exact statistical certainty, "the expert may give an opinion of a match to a

BY ADINA SCHWARTZ

reasonable degree of certainty in the ballistics field."5

As will be discussed, Judge Saris went on to exclude the particular testimony in *Monteiro* for failing to comport with the standards of documentation and peer review of the discipline of firearms and toolmark identification. While this aspect of her decision is highly useful to defense attorneys, the distinction Judge Saris made between statistical certainty and reasonable certainty within the field of firearms and toolmark identification is logically incoherent and has led to further bad case law.

In particular, in United States v. Diaz, Judge Alsup salved his conscience by holding that although the identifications in the case were admissible under Daubert, "[t]he experts may not, however, testify to their conclusions 'to the exclusion of all other firearms in the world.' They may only testify that a particular bullet or cartridge case was fired from a particular firearm to a 'reasonable degree of certainty in the ballistics field."6 Notwithstanding the quotation from Monteiro, Judge Alsup differed from Judge Saris in totally failing to understand the need for statistical empirical foundations for identity conclusions. Equating what firearms and toolmark examiners currently do with all that need be done, Judge Alsup wrote: "[I]t is impractical, even at this stage in the history of the art, to place a probability of error on individual analysis. The patterns do not lend themselves to probabilistic calculations. ... [T]he defense wants something that is beyond the state of the art. The holy grail of some in the profession is to find some way to impose upon firearms identification a statistical model to present to juries."7

There are two problems with restricting firearms and toolmark examiners to testifying that their identifications are based on "a reasonable degree of certainty in the ballistics field." First, this restriction is highly unlikely to make the jury aware that there are serious scientific problems with the discipline. Indeed, the government has responded to defense challenges by proposing that firearms and toolmark examiners testify that their conclusions have been reached to a reasonable degree of ballistic or scientific certainty.8

Second, as the Supreme Court recognized in *Kumho Tire*, "an expert's testimony [cannot be] reliable where the discipline itself lacks reliability, as, for example, do theories grounded in any so-called generally accepted principles of astrology or necromancy." Similarly,

due to the inherently probabilistic nature of identity conclusions, the absence of statistical empirical foundations makes the discipline of firearms and toolmark identification *itself* unreliable. Judge Saris to the contrary, it makes no sense to speak of reasonably certain conclusions within the field of firearms and toolmark identification as long as the requisite statistical empirical foundations do not exist.

Commendably, in 2008 in *United States v. Brown* and *United States v. Glynn*, Judge Rakoff prohibited firearms and toolmark examiners from testifying that their conclusions were reached to a reasonable degree of ballistics certainty or that they were based on science. In *Glynn*, Judge Rakoff wrote, "[W]hatever else ballistics identification analysis [can] be called, it [can] not fairly be called science. ... [I]ts methodology [is] too subjective to permit opinions to be stated to 'a reasonable degree of ballistic certainty." 10

Regrettably, Judge Rakoff went on to rule that even though it was not science and did meet the *Daubert* standard, firearms and toolmark identification testimony could still be admitted under Kumho Tire. The judge recognized that even under Kumho Tire, the admission of this testimony was problematic. "[B]allistics examination not only lacks the rigor of science but suffers from greater uncertainty than many other kinds of forensic evidence. ... The problem is how to admit it into evidence without giving the jury the impression — always a risk where forensic evidence is concerned — that it has greater reliability than its imperfect methodology permits."11 Judge Rakoff's solution was to restrict firearms and toolmark examiners to testifying that their identifications were more likely than not.

The qualifications that Judge Rakoff placed on the admission of firearms and toolmark identification testimony count as progress for the defense. Indeed, he recognized that since these identifications are only more likely than not, they cannot be the sole basis for criminal convictions. 12 At the same time, however, Judge Rakoff's decision to admit firearms and toolmark identification testimony with restrictions, rather than totally excluding such testimony, rests on the backwards assumption that if expert testimony is less intellectually rigorous than science, it is entitled to be judged by a less rigorous admissibility standard. To the contrary, as the Advisory Committee recognized in its Note to the 2000 amendment to Rule

702: "While the relevant factors for determining reliability will vary from expertise to expertise, the amendment rejects the premise that an expert's testimony should be treated more permissively simply because it is outside the realm of science. An opinion from an expert who is not a scientist should receive the same degree of scrutiny for reliability as an opinion from an expert who purports to be a scientist." ¹³

In addition, restricting examiners to testifying that their conclusions are more likely than not does not adequately take account of the statistical empirical problems with firearms and toolmark identification. As Judge Saris recognized in *Monteiro*, firearms and toolmark examiners "may not assert *any degree of statistical certainty, 100 percent or otherwise*, as to a match." Testimony that firearms and toolmark identifications are more likely than not is tantamount to testimony that these identifications have a certain degree of statistical certainty.

Defeating the Black Box Defense: Examiners Do Not Know It When They See It

Defense attorneys also need to counter the black box defense that regardless of how they do it, firearms and toolmark examiners reach accurate conclusions. Indeed, firearms and toolmark examiner Ronald Nichols has explicitly stated that "[w]hile oft criticized, the concept of 'I know a match when I see it' has its basis in [firearms and toolmark examiners' extensive] training."15 Nichols, an examiner at the Bureau of Alcohol, Tobacco and Firearms, has admitted, however, that "it's not surprising" and "not necessarily unexpected" for examiners to disagree about whether an inconclusive or an identification is the proper conclusion in a particular case.¹⁶ He and his colleagues have acknowledged that there is no standard for determining who is right in these situations and hence it is impossible to determine an actual error rate for casework.¹⁷ There is something basically wrong with a discipline that claims that its practitioners are trained to know it when they see it, even though different practitioners know mutually exclusive things when they see it and the discipline itself lacks resources to determine who is right.

Nichols has attempted to dismiss this criticism by claiming that when examiners disagree about the conclusion warranted in a particular case, the court should recognize that the discipline lacks standards for resolving such disputes and therefore "appropriately assign ... the task of weight to the jury."18 This purported solution places Nichols on the horns of a dilemma. On the one hand, if the jury is better qualified than examiners to decide whether identification conclusions are warranted in a particular case, the testimony of firearms and toolmark examiners fails to meet the most basic requirement for the admission of expert testimony: helpfulness to the trier of fact. On the other hand, if examiners are better qualified than the jury to reach identification conclusions, the jury cannot be qualified to resolve disagreements between examiners that the discipline itself lacks resources to resolve. In other words, the jury can only compensate for the fact that different firearms and toolmark examiners know different things when they see it if the testimony of examiners is unhelpful to the trier of fact and therefore not properly admitted as expert testimony.

Likewise, there is no reasonable basis for the compromise that Judge Gertner crafted in *Green* of not allowing the expert to testify that "the shell casings came from a specific ... pistol 'to the exclusion of every other firearm in the world,'" but allowing him to

"describe and explain the ways in which the earlier casings are similar to the shell casings test fired from the ... pistol found a year later." If the experts themselves are not qualified to assess the significance of the similarities they have observed between toolmarks, what justification can there be for assigning this task to a jury of non-experts?²⁰

Judge Gertner indicated that her decision to allow the examiner to testify as to his observations instead of excluding all firearms and toolmark identification testimony was primarily based on political, not scientific, considerations. "I reluctantly come to the above conclusion because of my confidence that any other decision will be rejected by appellate courts. ... While I recognize that the Daubert-Kumho standard does not require the illusory perfection of a television show (CSI, this wasn't), when liberty hangs in the balance — and, in the case of the defendants facing the death penalty, life itself — the standards should be higher than were met in this case, and than have been imposed across the country. The more courts admit this type of toolmark evidence without requiring documentation, proficiency testing, or evidence of reliability, the more sloppy practices will endure; we should require more."21

Inadequacy of the Existing Regime

The existing regime of education, laboratory accreditation, and proficiency testing does not compensate for either the absence of statistical empirical foundations or the impossibility of calculating an "actual" error rate for day-to-day work. Firearms and toolmark examiners need not - and often do not — have anything resembling a scientific education. For instance, the Massachusetts State Police officer who made the identifications in Monteiro did not have a college degree. Judge Saris ruled that he was qualified to testify as an expert witness, despite finding that his "scientific and academic credentials are underwhelming. He apparently has no formal scientific training, is neither certified by, nor is he a member of any professional organizations, [and] reads no literature in the field. ..."22 Similarly, although not mentioned in the Diaz opinion, the San Francisco Police Department officer who testified at the hearing in that case stated, in response to a question from the court about the probabilistic basis for identity conclusions, "I'm not a statistician. I'm not the greatest math wizard in the world."23



©2008 Hoftz Rubenstein Reminick LLP

Notwithstanding his admitted lack of scientific expertise, the officer presented a session entitled "Taking a Stand on *Daubert*! A *Daubert* Hearing in San Francisco — How It All Comes Together" at the annual AFTE (Association of Firearms and Toolmark Examiners) Training Seminar in 2007.²⁴

Indeed, one of prominent firearms and toolmark examiner Stephen Bunch's stated grounds for favoring the traditional, subjective approach Consecutive Matching Striae (CMS) or other attempts to develop statistical empirical foundations for identification conclusions is that firearms and toolmark examiners "may fail to understand or appreciate the research and the logic of interpreting this type of [statistical] evidence. Thus they may find it difficult to explain them to judge and jury. ... [This] could be a blow to the profession and to the administration of justice."25 In other words, according to Bunch, lack of scientific knowledge on the part of firearms and toolmark examiners is not the problem. Problems will arise only if an attempt is made to set identifications on scientific foundations; in that case, the "administration of justice" could be hindered because judges and juries might realize that examiners do not know whereof they speak.

The absence of stringent educational requirements for firearms and toolmark examiners is matched by the absence of any accreditation requirement for firearms and toolmark laboratories in the United States. Participation in the American Society of Crime Laboratory Directors (ASCLD) accreditation program is voluntary.26 In addition, even in an ASCLD-accredited laboratory, the proficiency of firearms and toolmark examiners is unlikely to be rigorously tested. Although the ASCLD requires each examiner in a laboratory to undergo yearly proficiency tests, except in the case of DNA laboratories, only one examiner per laboratory needs to take an external proficiency test with an approved ASCLD provider. The other examiners may take either internal or external tests.²⁷ Even if more than one examiner is tested by an approved ASCLD-provider, a laboratory need only authorize the provider to release the test results of one of its examiners to the ASCLD Laboratory Accreditation Board (LAB).28

Even firearms and toolmark examiners who are tested by an approved ASCLD-provider and have their results reported to the ASCLD/LAB do not have to undergo strenuous review. The only ASCLD-approved provider of external

proficiency tests for firearms and toolmark examiners is Collaborative Testing Services Inc. (CTS).²⁹ CTS itself cautions that the results on its tests "are not intended to be an overview of the quality of work performed in the profession and cannot be interpreted as such."³⁰

Major problems, which firearms and toolmark examiners themselves acknowledge, are that the CTS tests are declared, rather than blind, and present examiners with simpler problems than they encounter in actual casework. Ironically, in an article that purports to show that firearms and toolmark identification satisfies the *Daubert* standard, Grzybowski *et al.* admit that "[s]ince large numbers of tests have to be produced [by CTS] to uniform requirements, most tend to be rather straightforward and of only moderate difficulty." 32

To demonstrate that results on the CTS tests provide an inflated, rather than an accurate, estimate of the competence of examiners, the defense needs to take advantage of the online availability of the questions on the CTS tests as well as the scores, answers, and comments of the test takers. Even a quick perusal shows that in 2005 and 2006, the CTS firearms identification tests required examiners to eliminate bullets or cartridge cases that were fired from a gun of a different model and/or model and make than the one that fired the other sample bullets or cartridge cases. Examiners were not presented with the more difficult task of determining whether ammunition components were fired from the same or different guns of the same model and make.33

One examiner who took the 2006 CTS cartridge case test commented, "This test was straightforward and very easy. It took only a few minutes to make correct associations using toolmarks devoid of subclass influence. ... I suggest that you consider making the test more of a challenge in order to determine an error rate really reflective of actual casework where borderline cases are not uncommon."34 Similarly, test takers commented that one of the CTS cartridge case tests for 2005 was an "[e]asy test" and "[t]oo easy."35 Another examiner who took the test explained that the breech face marks and firing pin impressions on the cartridge case that was not fired from the suspect gun were "nothing like" those on the cartridge cases fired from the suspect gun.³⁶ On the other CTS firearms identification test administered in 2005, a test taker commented: "This was very easy. The class features of the firing pin in the firearm used to discharge items one and three [the cartridge cases not fired from the suspect gun] was [sic] so different it could be eliminated virtually with the naked eye."³⁷

Even though examiners were not required to distinguish between toolmarks produced by different guns of the same type, on each of the tests for 2005 and 2006, significant numbers of examiners were unable to determine which ammunition components were fired from the suspect gun. Two percent of those who took the 2006 bullet identification test (three examiners) wrongly identified the suspect gun as the source of one of the bullets it did not fire. The gun was misidentified as the source of another bullet by one percent of the test takers (two examiners), and 24 percent of the test takers (38 examiners) reached inconclusives, rather than eliminations, with regard to both of the bullets not fired by the suspect gun.³⁸ Inconclusives, rather than eliminations of cartridge cases that were not fired by the suspect gun, were reached by seven percent of test takers (19 examiners) on the 2006 CTS cartridge case test, and, in 2005, by five percent of test takers (12 examiners) on one of the CTS tests and four percent of test takers (five examiners) on the other CTS test.39 Additionally, on both the 2006 cartridge case test and one of the cartridge case tests for 2005, one percent (two examiners) reported inconclusives, rather than identifications, in regard to the cartridge case that was in fact fired from the suspect gun.40

By contrast to the 2005 and 2006 exams, the CTS bullet and cartridge case exams for 2007 required examiners to distinguish between toolmarks produced by different guns of the same model and make.41 On the 2007 bullet test, one percent of the test takers (two examiners) reached misidentifications; another 40 percent (114 examiners) reached inconclusives rather than eliminations in regard to the two bullets not fired from the suspect gun. One percent of the test takers reached inconclusives instead of identifications for the bullet that was fired by the suspect gun.42 On the 2007 cartridge case exam, five percent of the test takers (nine examiners) reached inconclusives rather than eliminations for the two cartridge cases not fired from the suspect gun. One percent (one examiner) reached an inconclusive, rather than an identification, in regard to the case that the suspect gun fired.43

Relatively low error rates are sometimes calculated for the CTS tests; for example, an error rate, consisting of both missed and misidentifications, of 1.4 percent and a misidentification rate of 0.6 percent on the CTS firearms tests from 1978 through 1991; a misidentification

rate of 1.3 percent on the CTS firearms tests from 1978-2002.44 These rates are determined, however, by not scoring errors when examiners reach inconclusive conclusions in regard to toolmarks that were in fact made by the same or different tools.45 As exemplified by the fact that 40 percent of the test takers reached inconclusives rather than eliminations when the 2007 CTS test asked them to distinguish between bullets fired from different guns of the same model and make, not scoring inconclusives as errors overestimates the competence of examiners who never report exclusions when class characteristics match.46 In addition, since the CTS tests are declared, rather than blind, eliminating inconclusives from the calculation of error rates overestimates the competence of examiners who react to their awareness of being tested by reporting inconclusives in situations where they would otherwise reach identifications.⁴⁷

Peterson and Markham's classic study counted inconclusives as errors, and found error rates, comprised of missed plus misidentifications, of 12 percent on the CTS firearms identification tests from 1978-1991 and 26 percent on the CTS toolmark identification tests from 1981-1991.48 Even these results are likely to understate actual, day-to-day error rates. Peterson and Markham found that "based on the number of tests and the hours of effort reported by laboratories on several tests ... many laboratories invested more time examining samples than would be expected or required on actual casework."49 This finding is consistent with Janine Arvizu's argument for blind, rather than declared, proficiency testing. "Although forensic analysts [in the tests in the Peterson and Markham study and other 'open' tests] do not know the 'true value' for a given proficiency sample, they are aware of the fact that a given sample is being used to assess their proficiency. Studies have shown that laboratory performance on this type of 'open' proficiency program is consistently better than on a program where the identification of proficiency samples is blind to the laboratory."50

The Peterson and Markham study is also likely to have underestimated day-to-day error rates because participation in the testing was voluntary, with about two-thirds of U.S. laboratories subscribing to the program, and one-third responding with data.⁵¹ The need to remove such "self-selection" bias and "survivorship bias" from proficiency testing has been recognized by firearms and toolmark examiners themselves.⁵²

On the basis of testimony about the problems with the CTS tests, Judge Rakoff

concluded in *Brown* that "there is no known error rate in any well-developed sense" for firearms and toolmark identification.⁵³ Despite recognizing that the absence of an error rate might be a reason for excluding scientific expert testimony under *Daubert*, Judge Rakoff reasoned that the firearms and toolmark evidence in the case might be admissible because "error rate is not irrelevant but it is certainly not a *sine qua non* for non-scientific expert testimony."⁵⁴

Law Enforcement and Confirmation Bias

Law enforcement and confirmation bias permeates both the day-to-day practice and the research of firearms and toolmark examiners. The traditional role of firearms and toolmark examiners is limited to "verifying investigative information."55 Examiners are only asked — and know they are only asked — to make test toolmarks with a gun or other tool and compare them to marks on ammunition components or other objects recovered from a crime scene when investigators have already linked the gun or other tool to a crime. Inherent in this situation is a danger of double counting. Instead of relying solely on resemblances between test and evidence toolmarks to reach identifications, examiners are also likely to be influenced, even if only unconsciously, by their knowledge that the gun or other tool has already been linked to a crime. The danger of double counting is heightened by the fact that virtually all firearms and toolmark laboratories in the United States are affiliated with law enforcement agencies. Many examiners, if not most, are initially trained as law enforcement officers.56

The bias in favor of confirming investigators' leads is likely to go unchecked because once a particular gun or other tool is identified as the unique source of evidence toolmarks, the standard practice of firearms and toolmark examiners is not to examine any other gun or tool to see if it might produce toolmarks that do at least as good a job at matching the evidence toolmarks. In addition, firearms and toolmark examiners tend to peer review each others' work only when an identification is reached.⁵⁷ One firearms and toolmark examiner has recognized that this limitation on the occasions for peer review is itself a source of confirmation bias. "[I]f the expert doing the check only ever checks positive matches, then his perception will be that whenever he sits at the microscope to conduct a peer review of casework, he will expect to see a positive match!"58

In sum, Judge Gertner's description of the examination in Green is true of many, if not most, firearms and toolmark examinations conducted in the United States. "The only weapon [the examiner] was shown was the suspect one; the only inquiry was whether the shell casings found earlier matched it. It was, in effect, an evidentiary 'show-up,' not what scientists would regard as a 'blind' test. [The examiner] was not asked to try to match the casings to the other test-fired Hi Point weapons in police custody, or any other gun for that matter, an examination more equivalent to an evidentiary 'line-up.' His work was reviewed by another officer, who did the same thing — checked his conclusions under the same conditions another evidentiary 'show-up."59 In accord with this, examiners Evan Thompson and Jan De Kinder state: "Rather than exclude a particular firearm or tool, most firearm/toolmark examiners probably find it far easier to include by reporting that 'this bullet was fired from a particular firearm' or 'this screwdriver was responsible for the striated toolmark left on the door."60

Similar bias infects the "studies" that the government's firearms and toolmark experts present in court. Some of these are "worst case studies" that purport to vindicate the hypothesis that each tool produces unique toolmarks by showing that even sequentially manufactured guns or other tools produce unique toolmarks. The vast majority of these studies are not blind, and it is hardly surprising that an examiner who sets out to vindicate the scientific foundations of his or her discipline should find significant differences among toolmarks that he or she knows were produced by different tools.

The government's firearms and toolmark experts are also likely to introduce "studies" that purport to show that the differences among the toolmarks produced by the same gun or other tool are so slight that the same gun or tool can be identified as the source of tens, or even hundreds, of marks. In the vast majority of these studies, however, the examiner knows that a single gun or other tool produced all the toolmarks he or she examines. Thus, the "studies" are inherently biased to vindicate the hypothesis that significant similarities persist among toolmarks produced by the same tool

Although the defense pointed out these methodological flaws, the judges in *Meeks and Warner* and *Diaz* invoked the government's proffered "studies" of the uniqueness and reproducibility of toolmarks as the basis for finding, respectively, that the *Daubert* factors of testability

and peer review and publication were satisfied.62 It is to be hoped that the NRC Report will put an end to such unwarranted judicial findings. In United States v. Brown, Judge Rakoff recognized that the report called into question firearms and toolmark identification's fundamental premises that each tool produces unique and reproducible marks. Two months earlier in United States v. Khalid Barnes, Southern District of New York Judge Stephen C. Robinson misguidedly insisted, however, that the NRC Report "does not identify any new evidence undermining the core premises upon which ballistics analysis is based, nor does it purport to."63 In United States v. English, Judge Geoffrey M. Alprin of the District of Columbia Superior Court similarly found that the NRC Report did not provide a basis for a Frye hearing.64

Excluding Particular Firearms and Toolmark Identifications

Here is the argument in the alternative: Even if firearms and toolmark identification testimony is not per se inadmissible, the particular identifications in the case should be excluded. Even judges who are presented with — and understand — detailed defense evidence of the systemic scientific problems are likely to be reluctant to exclude firearms and toolmark identification testimony across-the-board. As Judge Saris stated in Monteiro: "Courts have understandably been gun shy about questioning the reliability of firearm identification evidence."65 Epitomizing this "gun shyness," Judge Anita Brody of the federal district court for the Eastern District of Pennsylvania explained that because "there's rarely a case of any magnitude in ballistics or in arson or anything else that I don't get some of this testimony," she had become "agitated" when the defense began to develop an across-the-board challenge to the reliability and admissibility of firearms and toolmark identification in a bolt cutter identification case in 2004.66 Judge Brody stated, "What's concerning me is that this is a generic issue and I don't know whether the government recognizes it. ... I've been a judge for 23 years, nobody has ever challenged this. This is an issue that has great moment for the Department of Justice. ... If I preclude this testimony, it will make ripples all over the country."67

As in *Monteiro*, however, courts that are too "gun shy" to exclude firearms and toolmark identification testimony

across-the-board may nonetheless be willing to exclude the particular testimony in the case. As Judge Saris recognized, courts "must evaluate the reliability of not only the general field of toolmark identification but also the application [by the particular examiner]."68

The Absence of Documentation

Judge Saris relied on the absence of documentation to exclude the testimony of the Massachusetts State Police firearms examiner in *Monteiro*. There, the examiner's laboratory reports were confined to the statement that there was a "positive ID." He had not taken any photographs or made any sketches of the toolmarks on which his identifications were based nor made any notes about the basis for his identifications.

The absence of documentation was also a factor in the Florida Supreme Court's holding in 2001 in Ramirez v. State (Ramirez III) that the Frye standard was violated when firearms and toolmark examiners were allowed to identify the defendant's knife as the one and only one knife in the world that could have stabbed the victim. The court criticized the testimony of prosecution experts that "the examining technician generally takes no photomicrographs ... because lay persons would not be able to understand the identification process" and that "they [did] not prepare notes or written reports delineating the basis for identifications because to do so would not be helpful."69

A defense request for laboratory reports and bench notes is likely to be met by a bare bones report of an examiner's conclusions. Occasionally, an identification conclusion will be supplemented by the statement that test and evidence bullets matched on "fine striae" or that a cartridge case match was based on breech face marks, firing pin impressions, or some other type of mark. It is rare, however, for there to be photographs or sketches of the alleged matching toolmarks or descriptions of the resemblances that warranted an identification.

Following *Monteiro* and *Ramirez III*, the defense should argue that such an absence of documentation means that the identifications in the case should be excluded for failing to comport with the standards of the field of firearms and toolmark identification.⁷⁰ This argument should be supported by citations to the statements of firearms and toolmark examiners. The AFTE Theory of Identification, which government

experts tend to invoke to show that their field is scientific, states that "[t]he examiner is encouraged to report the objective observations that support the findings of toolmark examinations."71 Gryzbowski et al. recommend that each firearms and toolmark examiner "graphically demonstrate ... the basis for the opinion with the use of photographs" and that "comprehensive notes [be] taken that fully support the conclusions in the laboratory report."72 In an article quoted at length in Monteiro, firearms and toolmark examiner Bruce Moran argues strenuously for the use of photomicrographs, stating that "for our work to be valid, it must be verifiable to other examiners. ... Therefore, the data that we gather should provide a well-defined 'roadmap' as to what experiments we performed to answer the questions posed, what data was gathered, and a clear demonstration of the evidence from which we supported our conclusion(s)."73

As more defense challenges are brought, laboratories may attempt to cure an initial lack of documentation by having their staff re-examine evidence and document the basis for identifications. As Judge Gertner recognized in Green, such ex post facto attempts at justification are inherently unreliable because of the severe danger of confirmation bias. "Consider the forensic scientist who takes poor notes during an examination and prepares a skimpy report, but then goes back to 'spruce them up' shortly before trial. Even assuming the most honest of intentions, that examiner is inviting errors to infiltrate his conclusions and his testimony. The error potential of the original skimpy report, which leaves much to be supplied from memory, facilitates the creation of testimony more consistent with assumptions and later-acquired expectations than would be the case with a more detailed and complete contemporaneous account."74

There is a real life basis for the fear that examiners will be motivated to confirm, rather than neutrally examine, an initial identification when they are asked to provide documentation on the eve of trial. The Boston Police Department examiner in Meeks and Warner had not documented his initial identifications with photographs or diagrams. Three years later on the eve of trial, when the assistant district attorney asked him to comply with the defense request for photomicrographs, the examiner found that he could not "locat[e] the markings I matched in the initial comparison I made on 8/8/2002,

which resulted in ... finding consistent markings on only one land of the bullet from the victim's body."75 He therefore test fired another bullet from the revolver in 2005, and concluded, with the help of the new supervisor of the department, that three lands and two grooves of the bullet were useful for comparison and that the initial identification was supported by matching individual characteristics on two lands and one groove. 76 Both the examiner and his supervisor responded to the question of how they could be confident of the identification when the basis for finding a match had changed by testifying that "it is not rare for two examiners to find different areas of significance on the same piece of evidence."77

The judge in *Meeks and Warner* acknowledged that there was "much force" to my testimony that the initial reports in the case were so conclusory as to be "non-reports" and that the work in the case contravened the requirement that "[o] ne examiner be able to look at another examiner's report and then replicate the examination." Despite this, he admitted all the identifications in the case. Defense attorneys should not be deterred, however, from arguing in future cases that *ex post facto* attempts cannot cure initial failures to document the basis for identifications.

Nor should the defense be satisfied if, as was the case with the San Francisco Police Department Laboratory at the time of the Daubert hearing in United States v. Diaz, photomicrographs are the only documentation that examiners provide to support identifications. An examiner needs to indicate in writing, at the time of the examination, just what photographed features led him or her to conclude that the resemblances between evidence and test toolmarks were so great that the toolmarks must have been produced by the same gun or other tool. Otherwise, examiners may subvert the requirement of documentation by testifying, as the San Francisco Police Department examiner did in Diaz, that their identifications are based on both resemblances in the photomicrographs and resemblances that they recall in their mind's eye.79

Identifications in the Absence of a Gun

The government often seeks to implicate a defendant in crimes with testimony that bullets or cartridge cases recovered from various crime scenes and/or the defendant's home or other possessions must all have been fired by

the same gun, even though no gun has been recovered from the defendant or a crime scene. Such identifications are particularly unreliable because the absence of a gun makes it difficult to eliminate the possibility of subclass characteristics.80 In 2002 in Sexton v. State, the Texas Court of Criminal Appeals held that identifications made in the absence of a gun are unreliable and inadmissible.81 This holding was implicitly endorsed by the Fifth Circuit in United States v. Hicks in 2004.82 However, the defense should be wary of relying on these decisions because the Sexton and Hicks courts respectively hold that when a gun is recovered and test fired, a 100 percent certain identification is possible and "the error rate of firearms comparison testimony is zero or near zero."83

Particularly Strong Evidence of Law Enforcement Bias

In some cases, there might be strong evidence of pro-prosecution and confirmation bias that the defense can use to argue identifications are especially unreliable and inadmissible. For instance, in a federal case on which I worked, a request for bench notes and laboratory reports yielded a letter from the supervisor of the Los Angeles Police Department Laboratory thanking a detective for "explaining the reasoning behind your investigative hunch [linking two murders]" and stating that "because we are so committed to the task of serving your needs," an officer was assigned, contrary to ordinary procedure, to search the DRUGFIRE data base for the case numbers of the murders.84 "WHAMO! It looks very promising!" the supervisor wrote.85 Although she informed the detective that the identification would need to be confirmed by an examination of the actual ammunition, the conclusion of the letter strongly suggested that confirmation was a foregone conclusion. "Now, you need to take a moment to pat yourself on the back! Good guess! Good for you! Good police work!"86

Since the author of the letter supervised all the LAPD analysts in the case, the strong pro-prosecution bias evident in her letter cast severe doubt *both* on the identifications to which the letter referred *and* on the numerous other identifications in the case. The judge's refusal to grant an admissibility hearing epitomizes the fact that despite the severe scientific problems with firearms and toolmark identification, convincing judges to exclude this evidence remains uphill work.

Notes

- 1. Committee to Assess the Feasibility, Accuracy, and Technical Capability of a National Ballistics Database, National Research Council, *Ballistic Imaging* at 82 (National Academies Press 2008, *available at* http://books.nap.edu/catalog/12162.html ("NRC Report").
- 2. *Id.* at 82. *See also id.* at 85 (stating that firearms and toolmark examiners "should not overreach to make extreme probability statements").
- 3. United States v. Monteiro, 407 F. Supp. 2d 351, 373 (D. Mass. 2006); Adina Schwartz, A Systemic Challenge to the Reliability and Admissibility of Firearms and Toolmark Identification, 6 COLUM. SCI. & TECH. L. REV. 1 (March 28, 2005), available at http://www.stlr.org/cite.cgi?volume=6&article=2.
- 4. *Monteiro*, 407 F. Supp. 2d 351, 373 (emphasis added).
 - 5. Id. at 355 (emphasis added).
- 6. United States v. Diaz, Slip Copy, NO. CR 05-00167 WHA, 2007 WL 485967 (N.D. Cal. Feb 12, 2007), at 14.
- 7. Id. at 13. See also Commonwealth of Massachusetts v. Meeks and Warner, NO. CRIM.A. 2002-10961, CRIM.A. 2003-10575, 2006 WL 2819423 (Mass. Super. Sept. 28, 2006) at 31 ("This court finds that a database is neither practical nor necessary.").
- 8. See, e.g., Statement of Stephen G. Bunch, chief of the FBI Firearms and Toolmarks Unit, United States v. Worsley, 2003 FEL 6856, at 5 (Superior Ct. of the District of Columbia May 29, 2008) (stating that one "way to properly qualify an identification is to state that the examiner has matched a toolmark to a particular firearm to 'a reasonable degree of scientific certainty"").
- 9. *Kumho Tire Co. v. Carmichael*, 526 U.S. 137.151 (1999).
- 10. United States v. Glynn, No. 06 Cr. 580, 2008 WL 4293317, at *3 (SDNY Sept. 22, 2008). See also Transcript, United States v. Brown, No. 05 Cr. 538, at 11-14 (SDNY June 9, 2008); Transcript, United States v. Brown, at 1476 (SDNY June 18, 2008).
 - 11. Glynn at *6.
 - 12. Glynn at *6 n.14.
- 13. Advisory Committee, Note to the Amendment to Federal Rule of Evidence 702 (2000).
- 14. *Monteiro*, 407 F. Supp. 2d 351, 373 (emphasis added).
- 15. See Ronald G. Nichols, Defending the Scientific Foundations of the Firearms and Toolmark Identification Discipline: Responding to Recent Challenges, 52(3) J. FORENS. Sci. 586, 589 (May 2007).
- 16. Transcript of hearing, *United States* v. *Diaz* (N.D. Cal. January 24, 2007), at 51; Ronald Nichols, *The Scientific Foundations of Firearms and Toolmark Examination A Response to Recent Challenges*, CACNews 8,

26 (2nd Quarter 2006), available at http://www.cacnews.org/pdfs/2ndq06.pdf; Richard Grzybowski, Jerry Miller, Bruce Moran, John Murdock, Ron Nichols & Robert Thompson, Firearm Toolmark Identification: Passing the Reliability Test Under Federal and State Evidentiary Standards, 35 (2) Association of Firearms and Toolmark Examiners Journal ("AFTE J.") 209, 219 (Spring 2003).

17. Transcript of hearing, *United States v. Diaz* (N.D. Cal. January 24, 2007), at 53-54; Grzybowski, Miller, Moran, Murdock, Nichols & Thompson, *supra* n.16, at 219. *See also United States v. Diaz*, 2007 WL 485967 at 10 (erroneously reasoning that the *Frye* component of *Daubert* is satisfied despite "[t]he few critiques — such as the impossibility of calculating a true error rate and the fact that there can be no statistical, objective verification of an examiner's conclusions").

18. Ronald Nichols, *The Scientific Foundations of Firearms and Toolmark Examination* — *A Response to Recent Challenges*, CACNews 8, 26 (2nd Quarter 2006), *available at* www.cacnews.org/pdfs/2ndq06.pdf.

19. *United States v. Green,* 405 F. Supp. 2d 104, 109 (D. Mass. 2005).

20. See Joan Griffin and David LaMagna, Daubert Challenges to Forensic Evidence: Ballistics Next on the Firing Line, THE CHAMPION, September/October 2002, at 61-62 (a more extended argument against allowing firearms and toolmark examiners to testify about observed similarities and differences between toolmarks, but not allowing them to testify to identification conclusions); cf. Commonwealth of Massachusetts v. Meeks and Warner, NO. CRIM.A. 2002-10961, CRIM.A. 2003-10575, 2006 WL 2819423 at 47 (Mass. Super. Sept. 28, 2006) (allowing prosecution experts to testify as to both their observations of toolmarks and their identification conclusions on the ground that "ultimately [Judge Gertner's] distinction is artificial and unworkable. A jury, like a party involved with or affected by the real world of a scientific or technical field, is entitled to the judgment and opinion of an expert in that field as to the critical questions....").

- 21. *Green,* 405 F. Supp. 2d at 109 (footnotes omitted).
 - 22. Monteiro, 407 F. Supp. 2d at 373.
- 23. Transcript of hearing in *United States v. Diaz* (N.D. Cal. November 27, 2007) at 74.
- 24. Association of Firearms and Toolmark Examiners, *AFTE 2007-Tuesday, May 29, available at* http://www.afte.org/TrainingSeminar/AFTE2

007/Summaries/afte2007_tues.htm.

25. Stephen G. Bunch, Consecutive Matching Striation Criteria: A General Critique, 45(5) J. FORENS. SCI. 960 (2000).

26. American Society of Crime Laboratory Directors/Laboratory Accreditation Board,

About ASCLD/Lab, available at http://www.ascld-lab.org/dual/aslabdualaboutascld lab.html (last updated 01/05/07); Kristen A. Tomasetti, Analysis of the Essential Aspects of Striated Toolmark Examination and the Methods for Identification, 34(3) AFTE J. 289, 295 (Summer 2002). ("Participation in [the ASCLD accreditation] program is voluntary but strongly recommended to achieve standing in the forensic and legal communities.").

27. American Society of Crime Laboratory Directors/Laboratory Accreditation Board, *Proficiency Review Program* at 4, *available at* http://www.ascld-lab.org/international/pdf/ProficiencyReviewProgram.pdf (2005).

28. Id. at 2.

29. ASCLD/LAB, Approved Proficiency Test Providers, available at http://www.ascld-lab.org/international/aslabinternapprovedproviders.html (2006).

30.CTS, Firearms Examination Test No. 06-526 Summary Report at 1, available at http://www.collaborativetesting.com/reports /2626_web.pdf (2006).

31. See, e.g., Alfred Biasotti, John Murdock & Bruce Moran, Scientific Issues, in 4 David L. Faigman, Et al., Modern Scientific Evidence 544, 563-64 (2006-2007); Joseph L. Peterson and Penelope N. Markham, Crime Laboratory Proficiency Testing Results, 1978-1991, I, 40 J. Forens. Sci. 994, 997 (1995).

32. Grzybowski, Miller, Moran, Murdock, Nichols & Thompson, *supra* n.16, at 219.

33. See CTS, Firearms Examination Test No. 06-526 Summary Report, supra n.30, at 2; Firearms Examination Test No. 06-527 Summary Report at 2, available at http://www.collaborativetesting.com/reports/262 7_web.pdf; Firearms Examination Test No. 05-526 Summary Report at 2, available at http://www.collaborativetesting.com/reports/2526_web.pdf; Firearms Examination Test No. 05-527 Summary Report at 2, available at http://www.collaborativetesting.com/reports/2527_Web.pdf.

34. Firearms Examination Test No. 06-526 Summary Report, supra n.30, at 42. But see id. at 40 (comment by another test taker that "[t]he difficulty of the test was appropriate").

35. Firearms Examination Test No. 05-526 Summary Report, supra n.33, at 33.

36. Id. at 32.

37. Firearms Examination Test No. 05-527 Summary Report, supra n.33, at 21. But see id. at 22 (another test taker comments: "Very challenging test!").

38. See Firearms Examination Test No. 06-527 Summary Report, supra n.33, at 2,6.

39. See Firearms Examination Test No. 06-526 Summary Report, supra n.30, at 2, 8; Firearms Examination Test No. 05-526 Summary Report, supra n.33, at 2,7; Firearms Examination Test No. 05-527 Summary Report, supra n.33, at 2,6.

40. See Firearms Examination Test No. 06-

526 Summary Report, supra n.30, at 3, 8; Firearms Examination Test No. 05-527 Summary Report, supra n.33, at 3, 6.

41. See Firearms Examination Test No. 07-526 Summary Report at 2, http://www.collaborativetesting.com/reports/2726 _Web.pdf (2007); cf. id. at 45 (test taker's comment that the exam was "[a] difficult exercise requiring much more time than expected."); id. (another test taker's comment that "the comparison was a little more demanding than previous tests generally have been"); Firearms Examination Test No. 07-527 Summary Report, at 2.

42. Id. at 2, 8.

43. See Firearms Examination Test No. 07-527 Summary Report, supra n.41, at 2,7.

44. See Grzybowski et al., supra n.16, at 216; Nichols, Defending the Scientific Foundations, supra n.15, at 592.

45.Id.

46. See, e.g., Firearms Examination Test No. 07-526 Summary Report, supra n.41, at 45 (test taker's comment that "[t]his lab does not eliminate based on individual characteristics"); id. at 44-46 (similar comments by six other test takers); Evan Thompson & Jan De Kinder, Range of Exclusions, 38(1) AFTE J. 51, 51-52 (Winter 2006).

47. See Biasotti, Murdock & Moran, supra n.31, at 563 (stating that "some examiners may be more conservative when reporting the results of a declared proficiency test, feeling that they have little to gain but much to lose if they make an error"); Grzybowski et al., supra n.16, at 218.

48. Joseph L. Peterson & Penelope Markham, *Crime Laboratory Proficiency Testing Results*, 1978-1991, II, 40 J. FORENS. SCI. 1009, 1010, 1019, 1024 (1995).

49. Joseph L. Peterson & Penelope Markham, *Crime Laboratory Proficiency Testing Results*, 1978-1991, I, 40 J. FORENS. SCI. 994, 997 (1995) ("Peterson & Markham I").

50. Janine Arvizu, Forensic Labs: Shattering the Myth, THE CHAMPION, May 2000, at 18, available at http://www.nacdl.org/public.nsf/Champion Articles/2000may01.

- 51. Peterson & Markham I, supra n.49, at 997
- 52. See Stephen G. Bunch & Douglas P. Murphy, A Comprehensive Validity Study for the Forensic Examination of Cartridge Cases, 35(2) AFTE J. 201 (Spring 2003).
- 53. Transcript, *United States v. Brown,* at 1193 (June 13, 2008).

54. *Id.* at 1193-94. *See* text surrounding notes12-13, *supra*, for a discussion of why Judge Rakoff misinterpreted Federal Rule of Evidence 702 when he reasoned that because it was not scientific, the testimony of the firearms and toolmark examiner in the case was subject to a lower admissibility standard.

55. Richard E. Tontarski & Robert M.

Thompson, *Automated Firearms Evidence Comparison*, 43(3) J. FORENS. Sci. 641, 641-42 (1998).

56. For an excellent discussion of confirmation bias and other biases that infect the day-to-day work of forensic scientists, see D. Michael Risinger, Michael J. Saks, William C. Thompson & Robert Rosenthal, The Daubert/Kumho Implications of Observer Effects in Forensic Science: Hidden Problems of Expectation and Suggestion, 90 CALIF. L. REV. 1 (2002).

57. See United States v. Monteiro, 407 F. Supp. 2d 351, 369 (D. Mass. 2006) (reporting that the practice among firearms and toolmark examiners is to have matches peer reviewed); United States v. Diaz, Slip Copy, NO. CR 05-00167 WHA, 2007 WL 485967 (N.D. Cal. Feb 12, 2007) at 5 (failing to recognize the problems posed by the fact that "[t]he industry standard ... requires confirmation by at least one separate examiner when an identification is reached by the first examiner"); Commonwealth of Massachusetts v. Meeks and Warner, NO. CRIM.A. 2002-10961, CRIM.A. 2003-10575, 2006 WL 2819423 (Mass. Super. Sept. 28, 2006) ("Ideally, [peer review] would be 'blind' and done with respect to every examination. Budgetary restraints may well preclude that approach. ... [T]he Unit now does a peer review of match opinions. ... The Unit's procedures reasonably assure relia-

Boston, MA
August 5-8, 2009
Westin Copley Hotel

Bring a Revolution Home!
Trial Skills that Win

www.nacdl.org/meetings

bility; the shortcomings of the Unit's peer review procedures can be readily explored before the jury on cross-examination.").

58. Gerard Dutton, *Commentary: Ethics in Forensic Firearms Investigation*, 37(2) AFTE J. 79,82 (Spring 2005).

59. *United States v. Green,* 405 F. Supp. 2d 104, 107-08 (D. Mass. 2005).

60. Thompson & De Kinder, *supra* n.46, at 51.

61. See Diaz, 2007 WL 485967 at 6-8 (uncritically describing studies of the uniqueness and reproducibility of toolmarks); Commonwealth of Massachusetts v. Meeks and Warner, NO. CRIM.A. 2002-10961, CRIM.A. 2003-10575, 2006 WL 2819423 (Mass. Super. Sept. 28, 2006) at 43-44 (describing two nonblind studies of the reproducibility of toolmarks and a non-blind study of the uniqueness of toolmarks).

62. Meeks and Warner at 43-44 (relying on studies of the reproducibility of toolmarks, despite noting that they were not blind); Diaz at 6-7 (relying on studies of the uniqueness and reproducibility of toolmarks without acknowledging that it matters whether or not studies are blind).

63. Transcript, *United States v. Brown*, at 13 (June 9, 2008); *United States v. Khalid Barnes*, S9 04CR 186, slip opinion at 7 (SDNY April 2, 2008).

64. *United States v. Ronald English et al.*, Criminal N. 2007 CF1 16618 (Superior Ct. District of Columbia Criminal Division-Felony Branch March 20, 2008).

65.407 F. Supp. 2d at 364.

66. Transcript of hearing in *United States* v. Kain, Crim. No. 03-573-1 (E.D. Pa. February 24, 2004) at 101.

67.Id. at 87.Before the judge could rule on the *Daubert* challenge, the government dismissed the defendant's two-count indictment on a conspiracy charge and an arson charge that carried a five-year mandatory minimum sentence. The defendant pled guilty to one count of misprision of felony, and received a 10-month sentence, consisting of five months in a halfway house and five months in prison. *United States v. Kain*, Crim. No. 03-573-1 (E.D. Pa. March 15, 2004).

68.407 F. Supp. 2d at 357.

69. Ramirez v. State, 810 So.2d 836, 850-51 (Fla. 2001) ("Ramirez III").

70. *Monteiro*, 407 F. Supp. 2d at 355; *Ramirez III*, 810 So.2d at 851.

71. AFTE, Theory of Identification as It Relates to Toolmarks, 30(1) AFTE J. 86-88 (Winter 1998).

72. Grzybowski, Miller, Moran, Murdock, Nichols & Thompson, *supra* n.16, at 219.

73. Bruce Moran, *Photo Documentation of Toolmark Identifications* — *An Argument in Support*, 35 AFTE J. 174, 181 (2003), quoted in *Monteiro*, 407 F. Supp. 2d at 368.

74. 405 F. Supp. 2d at 121 (citation

omitted).

75. Affidavit of Adina Schwartz at 39 (emphasis added to examiner's letter to assistant district attorney); *Meeks and Warner*, 2006 WL 2819423 at 11.

76. Meeks and Warner, 2006 WL 2819423 at 12.

77.Id.

78. Id. at 33-34.

79. Transcript of hearing in *United States* v. *Diaz* (N.D. Cal. Nov. 27, 2006) at 135-37.

80. See the first part of this article in The Champion, October 2008, at 13. See also Alfred Biasotti, John Murdock & Bruce Moran, Scientific Issues, in 4 David L. Faigman, et al., Modern Scientific Evidence 544, 553 n.2 (2006-2007); CTS, Firearms Examination Test No. 03-526 Summary Report 35, available at http://www.collaborativetesting.com/reports/2326 web.pdf (2003).

81.93 S.W.3d 96, 101.

82.389 F.3d 514,525-26 (5th Cir. 2004).

83. Id. at 526; Sexton, 93 S.W.3d at 101.

84. Declaration of Adina Schwartz at 23-24, *United States v. Saldana*, CR 04-415 (A)-PA (C.D. Cal. Apr. 6, 2006). The original letter is available to defense attorneys on request.

85. *Id.* at 24.

86.Id.

© Adina Schwartz, 2008. All rights reserved.

About the Author

Adina Schwartz is a professor in the



Department of Law, Police Science and Criminal Justice Administration at John Jay College of Criminal Justice, and in the Criminal Justice Ph.D. Program, City University of

New York. Professor Schwartz has a Ph.D. in philosophy from the Rockefeller University and a J.D. from Yale Law School. She has served as an expert for the defense in 12 cases challenging the admissibility of firearms and toolmark identification. Her article, A Systemic Challenge to the Reliability and Admissibility of Firearms and Toolmark Identification, is cited in United States v. Monteiro and United States v. Green.

Adina Schwartz

John Jay College of Criminal Justice 899 Tenth Avenue New York, NY 10019 212-237-8402

Fax 212-237-8383

E-MAIL aschw81583@aol.com

THE CHAMPION