Challenging Firearms and Toolmark Identification — Part One

This article aims to convince defense attorneys that across-the-board challenges to the scientific stature and admissibility of firearms and toolmark identification should be brought in every case involving such evidence. In addition, the bench notes of the prosecution’s firearms and toolmark examiner and associated police reports should be obtained and carefully scrutinized in every case. Very likely, this evidence will enable the defense to argue that even if firearms and toolmark identifications are not per se inadmissible, the particular identifications in the case should be excluded because they do not comport with the standards of the field.

Recent changes in the legal landscape make defense challenges substantially easier. In an article published in the Columbia Science and Technology Law Review in spring 2005, I stated that “both before and after Daubert, firearms and toolmark identification testimony has largely been admitted as a matter of course. No court … has recognized the systemic scientific problems with the field.” After the article was published, two decisions in the federal district court for Massachusetts — United States v. Green, decided by Judge Gertner in late 2005, and United States v. Monteiro, decided by Judge Saris in early 2006 — took major steps towards recognizing the systemic scientific problems and excluding firearms and toolmark identifications. Although courts continued to admit firearms and toolmark identifications after those decisions, several departed from the longstanding judicial tendency to take admissibility for granted and addressed the scientific issues at length.

Most recently, in United States v. Glynn in September 2008, the defense scored a major victory. There, Judge Rakoff of the federal district court for the Southern District of New York recognized that “ballistics examination not only lacks the rigor of science but suffers from greater uncertainty than many other kinds of forensic evidence.” On this basis, the court concluded that to allow any “ballistics examiner … to testify that he had matched a bullet or casing to a particular gun ‘to a reasonable degree of ballistic certainty’ would seriously mislead the jury,” and held that “ballistics opinions may be stated in terms of ‘more likely than not,’ but nothing more.”

Also in 2008, a substantial boost to defense challenges was provided by a report by the National Research Council Committee to Assess the Feasibility, Accuracy, and Technical Capability of a National Ballistics Database (the NRC Report). The Committee’s mandate expressly precluded it from evaluating the admissibility of firearms identification evidence. Nonetheless, the Committee explicitly found that there are serious scientific problems with both the underlying premises of firearms and toolmark identification and the method firearms and toolmark examiners use to reach identifications.

Judge Saris’ words in Monteiro aptly describe the current legal situation: “Courts have understandably been gun shy about questioning the reliability of firearm identification evidence. Storm clouds, however, are gathering.” By analyzing the scientific issues, legal precedents, and relevant portions of the NRC Report, this article aims to pro-
vide defense attorneys with a blueprint for admissibility challenges and for countering firearms and toolmark identification testimony in the event that it is admitted.

What Is Firearms and Toolmark Identification?

Firearms identification is often confused with ballistics, which, properly speaking, deals with the motions that firearms impart to projectiles (i.e., bullets and shot). By contrast, the forensic science discipline of toolmark identification aims to identify particular tools, such as a gun barrel or a bolt cutter, as the unique source of marks on crime scene evidence, such as a bullet or a piece of fence. Toolmarks are a subspecies of toolmark identification dealing with the toolmarks that bullets, cartridge cases, and shotshell components acquire by being fired and that unfired cartridge cases and shotshells acquire by being worked through the action of a firearm. Although this article’s focus is on firearms identification, the criticisms developed here extend to toolmark identification as a whole.1

Firearms and toolmark examiners compare evidence toolmarks on ammunition components or other objects recovered from crime scenes with test toolmarks that they produce on other ammunition components or objects by firing or otherwise using a particular gun or other tool. If the same class characteristics are found on evidence and test toolmarks (for example, the same number of rifling impressions on a test-fired bullet and an evidence bullet recovered from a crime scene), an examiner uses a comparison microscope to compare the toolmarks’ individual characteristics (for example, microscopic lines within rifling impressions). The object is to determine whether the individual characteristics are so similar that one and the same tool (for example, a particular gun barrel) must have produced both the test and the evidence toolmarks.

Once a particular gun or other tool is identified as the source of evidence toolmarks, the practice of firearms and toolmark examiners is not to examine any other gun or tool to see if it could produce toolmarks that would do as good or better a job at matching the evidence toolmarks. As Judge Gertner recognized in Green: “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.”20

How Can Firearms and Toolmark Identification Be Challenged?

The premise underlying firearms and toolmark identification is that a tool, such as a firearm barrel, leaves a unique toolmark on an object, such as a bullet. An equally crucial premise is that toolmarks are reproducible. As the NRC Report recognized: “To be useful for identification, the characteristic marks left by firearms must not only be unique but reproducible — that is, the unique characteristics must be capable of being deposited over the multiple firings so that they can be found on recovered evidence and successfully compared with those on other items.”21

The NRC Committee found that the basic premises of firearms and toolmark identification had not been scientifically established. The NRC Report stated that “the validity of the fundamental assumptions of uniqueness and reproducibility of firearm-related toolmarks has not yet been fully demonstrated.”22 According to the NRC Report, extensive, basic scientific research is needed to determine whether the assumptions are valid.

Additional general research on the uniqueness and reproducibility of firearm-related toolmarks would have to be done if the basic premises of firearms identification are to be put on a more solid scientific footing.

Fully assessing the assumptions underlying firearms identification would require careful attention to statistical experimental design issues, as well as intensive work on the underlying physics, engineering and metallurgy of firearms, but is essential to the long-term viability of this type of forensic evidence.23

The prestige of those voicing skepticism makes the report’s doubts about the basic premises of firearms and toolmark enormously useful for admissibility challenges. The National Research Council is the operating agency of the National Academy of Sciences, an independent body of distinguished scientists that Congress established in 1863 for the purpose of advising federal government agencies on scientific and technical questions.24 NRC committees are staffed by top scientists and professionals who work on a voluntary basis.25 The appointment process is designed to ensure that committee members have an “appropriate range of expertise for the task” and bring “a balance of perspectives” to a project.26

In Glyn, Judge Rakoff invoked the NRC Report to rule that the assumptions that each gun produces unique and reproducible toolmarks had not been scientifically established.27 In United States v. Brown, a June 2008 case in which he had ruled from the bench, Judge Rakoff more forcefully recognized that the NRC Report raised fundamental questions about the reliability of firearms and toolmark identification.

Twice in that report in bold face so that no one can miss it, the authors of the report who appear to include quite a few notable scientists as well as others, state, “Finding: The validity of the fundamental assumptions of uniqueness and reproducibility of firearms-related toolmarks has not yet been fully demonstrated.

So, that goes to the most basic premise before we get into anything else, the most basic premise on which this, what you [the Assistant United States Attorney] call ballistic science is premised, yes?”28

Nonetheless, from both a legal and scientific point of view, the best strategy is to argue both that the NRC Report casts severe doubt on the underlying premises of firearms and toolmark identification and that even if guns are capable of leaving unique and reproducible marks, there is no scientific basis for examiners’ claim to be able to single out the particular gun that produced an evidence toolmark. This two-pronged strategy enables attorneys to build on the defense victories in Monteiro, Green, and Glyn. Although Judge Saris concluded in Monteiro that the defense attack on the uniqueness of toolmarks had “ fizzled at the hearing,”29 she went on to recognize that “[t]he question of whether the methodology of identifying a match between a particular cartridge case and gun is reliable requires far more analysis.”30

Similarly, Judge Gertner reasoned in Green that “even assuming that some of these marks are unique to the gun in question, the issue is their significance, how the examiner can distinguish one from another, which to discount and which to focus on, how qualified he is to do so, and how reliable his examination is.”31 Despite recognizing that the underlying premises of firearms and toolmark identification had not been scientifically
established, Judge Rakoff ruled, in Glynn, that the uniqueness of toolmarks “has been sufficiently well-documented as to support a reasonable hypothesis of its validity.”26 He also ruled that the reproducibility of toolmarks “is both plausible and sufficiently documented by experience as to provide a good working assumption for most practical purposes.”27 Nonetheless, Judge Rakoff went on to hold that because “ballistics comparisons involve the exercise of a considerable degree of subjective judgment” and “lack … defining standards to a degree that exceeds most other types of forensic expertise,” substantial restrictions must be placed on the testimony of firearms and toolmark examiners.28

The remainder of this article will focus on the problems with the method of firearms and toolmark identification.

Types of Toolmarks

To understand the scientific problems with the method firearms and toolmark examiners use to make identifications, it is necessary to distinguish between various types of toolmarks. Toolmarks are either striated toolmarks consisting of patterns of scratches or striae produced by the parallel motion of tools against objects (e.g., the marks the barrel of a gun produces on fired bullets) or impression toolmarks produced on objects by the perpendicular, pressurized impact of tools (e.g., breech face marks that the breech face of a gun produces on fired cartridge cases). Both types of toolmarks have class, subclass, and individual characteristics.

The distinctively designed features of tools are reflected in the class characteristics of the toolmarks produced by all tools of a certain type. For example, the rifling impressions on bullets are class characteristics reflecting the number, width, and direction of twist of the lands and grooves in the types of barrels that fired them. By contrast to class characteristics, microscopic individual characteristics are purported to be unique to the toolmarks each individual tool produces and to correspond to random imperfections or irregularities on tool surfaces produced by the manufacturing process and/or subsequent use, corrosion or damage.

Although firearms and toolmark examiners frequently state that every tool produces toolmarks with unique individual characteristics, this is, at best, sloppy speech. It is well known within the discipline that only some manufacturing processes result in firearms or other tools with such differentiated surfaces that each tool produces toolmarks with unique, individual characteristics. Other manufacturing processes create batches of tools, such as breech faces of firearms, with similarities in appearance, size, or surface finish that set them apart from other tools of the same type. The toolmarks produced by tools in the batch have matching microscopic characteristics, called subclass characteristics, which distinguish them from toolmarks produced by other tools of the type. In addition to the subclass characteristics that are common to all toolmarks produced by tools in the batch, individual characteristics may or may not also be present on the toolmarks produced by newly manufactured tools in the batch. While wear and tear on some tools may cause the subclass characteristics on their toolmarks to be completely replaced by individual characteristics, in other tools subclass characteristics may persist alongside individual characteristics.29

Central Pitfalls in Firearms And Toolmark Identification

Three central pitfalls stand in the way of firearms and toolmark examiners’ goal of identifying one and only one tool as the source of a particular toolmark.

The First Difficulty: Individual Characteristics of Toolmarks Are Combinations of Non-Unique Marks

A first barrier in the way of reliably identifying the source of an evidence toolmark tends to be obscured by firearms and toolmark examiners’ ambiguous use of the term “individual characteristics.” Examiners sometimes use the term to refer to the entire unique microscopic marks that are allegedly produced by individual tools. At other times, the term “individual characteristics” is used to refer to the component microscopic marks, not in themselves unique to any tool, that come together as a pattern to comprise the microscopic marks that are allegedly unique to particular tools. The component nature of the individual characteristics of toolmarks was recognized as early as 1935: “It is probably true that no two firearms with the same class characteristics will produce the same signature, but it is likewise true that each element of a firearm’s signature may be found in the signatures of other firearms. … An individual peculiarity of a firearm can, therefore, be established by elements of identity which form a combination the coexistence of which is highly improbable in the signature of other firearms with the same class characteristics.”30

As a result of the overlapping individual characteristics of the toolmarks made by different tools, misidentifications may result because examiners assume that a certain amount of resemblance proves that the same tool produced both test and evidence toolmarks, when the same amount of resemblance is possible between toolmarks produced by different tools. According to prominent firearms and toolmark examiners Alfred Biasotti, John Murdock, and Bruce Moran, many of the disagreements between examiners about the conclusions warranted in a particular case “stem from one examiner ascribing too much significance to a small amount of matching striae and not appreciating that such agreement is achievable in known non-match comparisons.”31

In the 1990s, the development of the Bureau of Alcohol, Tobacco and Firearms’ (BATF) computerized comparison system, IBIS (Integrative Ballistics Information System), made it more evident that misidentifications could result from examiners’ failure to appreciate how similar toolmarks produced by different tools could be. For the first time, IBIS enabled examiners to compare the toolmarks on vast numbers of bullets and cartridge cases. Their comparisons had previously been limited to toolmarks encountered in their casework and training.

Firearms and toolmark examiner Joseph J. Masson observed that as the number of toolmarks inputted into the IBIS database grew, toolmarks produced on bullets by different guns of the same caliber were found to be more similar than examiners had previously believed they could be.32 The extent of the possible similarities between toolmarks produced by different tools was also highlighted by studies finding that as the IBIS database was expanded to include increasing numbers of cartridge cases that had been test fired by guns of the same caliber and make, the top 10 or even 15 candidate matches that IBIS listed for a queried cartridge case increasingly did not include the cartridge case known to have been fired by the same gun.33 The NRC Report stated that these studies “compellingly demonstrate that [IBIS’s] performance can degrade in databases flooded with same-class-characteristic images.”34

The Second Difficulty: Individual Characteristics of Toolmarks Change Over Time

A further barrier in the way of reliable firearms and toolmark identifications is that the individual characteristics of the marks a particular tool makes change over
time. Indeed, firearms and toolmark examiners do not expect the toolmarks on bullets fired from the same gun to ever be exactly alike. The changes in toolmarks reflect the changes in a tool’s surfaces that occur as the tool is used, and/or as damage or corrosion occurs. An additional cause of differences among the toolmarks a particular gun leaves on ammunition is that pressures and velocities involved in the physical interaction between the weapon and the ammunition at firing are subject to intrinsic variation from shot to shot, thus resulting in variations of the shape, orientation, and localization of the signature markings, even for the same combination of firearm/ammunition type. In Monteiro, Judge Saris recognized, “A perfect correspondence between the lines on a test fired cartridge and the evidence recovered from the scene is impossible; in the real world, there is no such thing as a ‘perfect match.’” Similarly, Judge Gertner stated in Green, “Just because the marks on the [cartridge] casings are different does not mean that they come from different guns. Repeated firings from the same weapon, particularly over a long period of time, could produce different marks as a result of wear or simply by accident.” In United States v. Glynn, Judge Rakoff recognized that an inherent difficulty for firearms and toolmark identification arises from the fact that “a gun barrel may itself change slightly with each firing, such that it may leave different impressions on a casing [sic] depending on when during the gun’s life a shot is fired. Casings from the same firearms may appear markedly different because of an irregular firing or because of the manner in which they hit against various materials.”

As a consequence of the impermanence of toolmarks, differences between evidence and test toolmarks will sometimes be correctly attributed to changes in the surfaces of the suspect tool between the time the evidence and test toolmarks were made. At other times, such an attribution will be wrong; the evidence and test toolmarks differ because the source of the evidence mark was a tool similar, but not identical, to the suspect tool. While misidentifications may accordingly occur if examiners wrongly attribute differences in test and evidence toolmarks to changes in the same tool over time, identifications may also be missed if examiners fail to realize that differences between test and evidence toolmarks are compatible with their having been produced by the same tool.

Ronald Nichols, a firearms and toolmark examiner with the BATF who is one of the most ardent defenders of the discipline, has attempted to dismiss this difficulty by suggesting that the impermanence of toolmarks can cause examiners to miss identifications, but not to make misidentifications. However, in the 1959 study that Nichols described in 1997 as “the most exhaustive statistical empirical study ever published,” Biasotti found matches of 21-38 percent and 15-20 percent of the striae per land or groove impression on bullets respectively fired from the same and different .38 Special Smith & Wesson revolvers. This near-complete overlap in the amount of similarity in toolmarks produced by the same and different guns strongly suggests that examiners can make misidentifications by wrongly attributing differences between toolmarks made by different tools to changes in the same tool over time.

The Third Difficulty: Danger of Confusing Subclass With Individual Characteristics

A third major barrier in the way of reliable firearms and toolmark identifications arises in regard to the interpretation of complete (or near total) similarity between evidence and test toolmarks. Instead of being an individual characteristic unique to the toolmarks produced by a single tool, the matching, complete microscopic pattern might be a subclass characteristic common to all toolmarks produced by tools in a batch. In Monteiro, Judge Saris cited my article in the Columbia Science and Technology Law Review and stated: “A recent article has highlighted the complexity of comparing patterns because of the difficulty in distinguishing between class, subclass and individual characteristics, noting that a firearm ‘may be wrongly identified as the source of a toolmark it did not produce if an examiner confuses subclass characteristics shared by more than one tool with individual characteristics unique to one and only one tool.” Likewise, in Green, Judge Gertner recognized that “[p]lainly, confusing individual characteristics with class or subclass ones could lead to false negatives, as well as false positives.”

In a defense of firearms and toolmark identification against me and other critics, Ronald Nichols cited 19 studies of subclass characteristics by firearms and toolmark examiners. At most, however, the cited studies provide rough rules of thumb about circumstances in which subclass characteristics are or are not likely to occur. They do not provide either strict rules for determining whether a microscopic pattern on a toolmark is an individual or a subclass characteristic or strict rules as to which tools or manufacturing processes do or do not produce toolmarks with subclass characteristics. To avoid misidentifications based on confusing subclass characteristics shared by more than one tool with individual characteristics unique to one and only one tool, examiners need to rely on personal familiarity with types of forming and finishing processes and their reflections in toolmarks.

In accord with this, Biasotti and Murdock explain that a risk of misidentifications arises because “some machining processes are capable of reproducing remarkably similar surface characteristics (i.e., gross contour and/or fine striae, etc.) on the working surfaces of many consecutively produced tools which if not recognized and properly evaluated could lead to a false identification.” They warn that “[t]he examiner must ... be familiar with the various forming and finishing processes in order to distinguish those ... surface characteristics that are truly individual from those surface characteristics that may characterize more than one tool.” Indeed, in replying to my Columbia Science and Technology Law Review article, Nichols himself acknowledged that the literature on subclass characteristics does not suffice to prevent examiners from confusing subclass with individual characteristics: “[T]here is not one conscientious firearms and toolmark examiner who would suggest that personal familiarity with tool finishing processes and their effects on tool surfaces is anything but vital to the proper understanding of subclass characteristics. Without such knowledge and appreciation of manufacturing techniques, examiners would have no way of ascertaining if subclass characteristics could exist.”

Contradicting his testimony in United States v. Diaz in January 2007 that “it’s not very difficult” to distinguish between subclass and individual characteristics, Nichols recently wrote that “[t]he difficulty of addressing subclass characteristics is not in debate.”

The danger that misidentifications will result from confusing subclass with individual characteristics is real, not theoretical. Prominent firearms and toolmark examiner Bruce Moran reported that this type of confusion produced misidentifications of striated toolmarks in the 1980s. In response, members of the Association of Firearms and Toolmark Examiners (AFTE) formed the Criteria for Identification Committee. The term “subclass characteristics” was coined in 1989 and incorporated in the AFTE glossary definitions in 1992. Despite prominent firearms and tool-
mark examiners’ warnings that “[c]autious should be exercised in distinguishing subclass characteristics from individual characteristics,” in my experience, many prosecution experts do not acknowledge the possibility of subclass characteristics in their testimony or bench notes. In Brown, Judge Rakoff ruled that because of his failure to take account of subclass characteristics, the prosecution expert would only be allowed to testify that his identifications were “more likely than not.” In reaching this ruling, the judge recognized that “many commentators had noted the real possibility that subclass characteristics could lead to false positives. ...” The judge was troubled by the evidence at the Daubert hearing “that both his [the prosecution expert’s] standard practice was pretty much to ignore them [subclass characteristics] and he had ignored them in this case or, that is to say, made no investigation as to whether they would have been present in this case, and he didn’t know how you could distinguish individual characteristics from subclass characteristics if they had existed in this case.”

The issue of whether subclass characteristics have been adequately distinguished from individual characteristics is particularly important when a gun is linked to a defendant through a cartridge case identification. Frequently, such identifications are based on “matching” breech face marks or firing pin impressions, as breech face marks and firing pin impressions are the only types of toolmarks that can be used to show that a cartridge case was fired, as opposed to cycled through a particular gun. Regardless of how it was manufactured, it is possible for the breech face of a gun to produce toolmarks with subclass characteristics. Nichols states: “Breech face marks can be cut, milled or stamped. In each case, subclass characteristics may be produced.” Similarly, on the basis of studies finding subclass, rather than individual, characteristics on firing pin impressions, prominent firearms and toolmark examiners have warned that reliable firearms identifications cannot be based on firing pin impressions alone.

In addition, distinguishing between subclass and individual characteristics is particularly difficult when a crime scene does not yield any gun whose class characteristics match those of the ammunition components recovered from the scene. In this situation, firearms and toolmark examiners sometimes compare various ammunition components recovered from the crime scene and/or other crime scenes or the suspect’s home or possessions and conclude that they all must have been fired (or worked through) the same gun. Biasotti, Murdock, and Moran acknowledge that because “[t]he most reliable way to assess the potential for subclass influence in a toolmark is by direct examination of the responsible tool working surface that produced the mark,” it is particularly difficult for examiners to make identifications in the absence of a gun. Similar reasoning has led some firearms and toolmark examiners to criticize the Collaborative Testing Service (CTS) proficiency tests for asking them to make identifications in the absence of a gun. On a 2003 test, one examiner commented that “[a] cast of the firearm’s breech face would have been taken to rule out any subclass characteristics from the similar ammunition used for tests in this comparison.” Another wrote: “In an actual case, I would not except [sic] test-fired cartridge cases from another agency or intraagency. I would want to examine the tool working surfaces of the firearm in order to eliminate the possibility of subclass carry over.”

The problems with identifications made in the absence of a gun are magnified by the proliferation of makes and models of guns (for example, over 2,000 different makes and models of semiautomatic handguns are sold in California each year). In one case on which I worked, the bench notes of examiners from the Los Angeles Police Department stated that guns from too many manufacturers to list could have produced toolmarks with the general rifling characteristics (i.e., class characteristics) that had been found on various spent ammunition components. Their inability to narrow down the types of guns that could have fired the ammunition components made it impossible for the LAPD examiners to use knowledge of how various types of guns are manufactured to determine whether the matching microscopic characteristics they found were subclass or individual characteristics. Nonetheless, the examiners concluded, without even considering the possibility of subclass characteristics, that the ammunition components must have been fired from the same gun.

Statistical Empirical Foundations

Firearms and toolmark identifications are not based on adequate statistical empirical foundations. Taken together, the three major pitfalls to reliable firearms and toolmark identifications imply that firearms and toolmark identifications are inherently probabilistic. On the one hand, shared subclass characteristics and/or similarities between the marks comprising the individual characteristics of toolmarks create substantial resemblances between toolmarks produced by different tools. On the other hand, even toolmarks produced by the same tool do not perfectly match. The similarities between toolmarks made by different tools and the differences between toolmarks made by the same tool imply that a statistical question must be answered to determine whether a particular tool was the source of a toolmark on an object recovered from a crime scene. What is the likelihood that the toolmarks made by a randomly selected tool of the same type would match — as closely as the toolmarks made by the suspect tool — the characteristics of the evidence toolmark?

Firearms and toolmark examiners implicitly deny that identity determinations are probabilistic when, as is typical, they testify that they have identified a firearm or other tool as the source of an evidence toolmark, to the exclusion of all other tools in the world. Ronald Nichols has advised examiners to still the quibbling of attorneys by abandoning absolute identity conclusions and testify that the likelihood that any tool besides the suspect tool could have produced the evidence toolmarks is so small that it can, for all practical purposes, be ignored. The problems with firearms and toolmark examiners’ testimony are not linguistic, however, but scientific: the requisite empirical and statistical foundations have not been laid for either absolute or probabilistic identification conclusions.

The Traditional Subjective Approach

The prevailing practice of firearms and toolmark examiners conforms to “the stereotype [of] the distinguished, greying individual on the stand saying, ‘my opinion is based on my many years of experience in the field.’” Each examiner’s identity conclusions are based solely on his or her own inarticulable, mind’s eye judgments of when the resemblances between toolmarks are so great that they must have come from the same tool. There is no attempt to articulate how many or what types of resemblances between toolmarks are necessary for identity conclusions, much less to develop relevant and representative data bases and calculate the frequencies with which various microscopic characteristics of toolmarks occur. As Biasotti recognized, when examiners follow this subjective approach, they implicitly
admit that “we lack necessary statistical data which would permit us to formulate precise criteria for distinguishing between identity and nonidentity with a reasonable degree of certainty.” Other examiners have also criticized the subjective approach for conflicting with the scientific value of “as far as possible, support[ing one’s] opinion by reference to logical reasoning and an established corpus of scientific knowledge.”

Although examiners are likely to invoke the AFTE Theory of Identification as proof that their discipline rests on firm scientific foundations, the AFTE Theory does nothing to cure the absence of statistical empirical foundations for firearms and toolmark identifications. The theory states that there is an exceedingly small likelihood that any tool besides the suspect tool produced the evidence toolmarks when the observed agreement between test and evidence toolmarks is superior to that of the best known nonmatch and consonant with that of the best known match. Nichols has acknowledged that “there is no universal agreement as to how much correspondence exceeds the best known nonmatching situation.” He fails to realize, however, that the absence of agreement implies that the AFTE Theory is vacuous; it provides examiners with no guidance about when to declare a match.

In contrast to Nichols, Judge Saris recognized in Monteiro that the AFTE Theory “leaves much to be desired. … [I]t is not a numeric or statistical standard, but is based on the individual examiner’s expertise.” On this basis, she criticized the theory for being “tautological: it requires each examiner to decide when there is ‘sufficient agreement’ of toolmarks to constitute an ‘identification.’” Similarly, Judge Rakoff reasoned in Glyn that “ballistics opinions are significantly subjective [because] the standard defining when an examiner should declare a match – namely, ‘sufficient agreement’ – is inherently vague.”

The CMS Identification Criterion

In 1997, Biasotti and Murdock proposed the Consecutive Matching Striae (CMS) criterion for the identification of striated toolmarks. Under CMS, the threshold for identifying a particular tool as the source of a three-dimensional toolmark is a match between evidence and test toolmarks of one group of six consecutive matching striae or two different groups of at least three consecutive matching striae in the same relative position. The threshold for two-dimensional toolmarks is one group of eight consecutive matching striae or two groups of at least five consecutive matching striae in the same relative position.”

The development of CMS was motivated by Biasotti’s criticisms of the traditional subjective approach. Starting in the 1950s, Biasotti criticized the “almost complete lack of factual and statistical data pertaining to the problem of establishing identity in the field of firearms identification. …” He wrote that “[i]f we accept the present apparent state of development as adequate and believe that no objective statistical data for establishing identity can be developed, then the subject of firearms and toolmark identification will remain essentially an art limited by the intuitive ability of individual practitioners.”

Perceived as an alternative to the traditional approach, CMS encountered widespread resistance on the part of firearms and toolmark examiners. According to Stephen G. Bunch, a firearms and toolmark examiner who is one of the most prominent critics of CMS: “Since A.I. Biasotti conducted his original identification-criteria research in the 1950s, [there has been a] debate over the relative virtues of objective and subjective methods in forensic firearms identification — specifically over the virtues of counting consecutive matching striations on bullets. …”

In an attempt to downplay the controversy over the relative merits of CMS and the traditional subjective approach (and to defuse the claim that firearms and toolmark identification does not satisfy the Frye standard), Nichols has insisted that “CMS is not a more objective way of performing examinations but simply a means by which an examiner can describe what he or she is observing in a striated toolmark comparison.” At the same time, Nichols has described CMS as an attempt “to standardize the concept of the best-known nonmatch discipline-wide.”

These two descriptions of CMS cannot both be true, given Nichols’s own admission, in both testimony and publications, that under the traditional approach, “difference[s] between examiners as to what constitutes the best-known nonmatch situation” make it “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. On the other hand, if CMS is not “a different method than has been practiced throughout the years,” the CMS identification criterion must be such a malleable...

THE NACDL INDIGENT DEFENSE COMMITTEE INVITES NOMINATIONS FOR THE 2009 Champion of Indigent Defense Award

The NACDL Champion of Indigent Defense Award recognizes an individual for exceptional efforts in making positive changes to a local, county, state, or national indigent defense system. Although the outstanding representation of every indigent defendant is one of NACDL’s foremost goals, this award is intended to highlight efforts toward positive systemic changes through legislation, litigation or other methods and not the outstanding representation of individual clients.

The Champion of Indigent Defense Award is awarded annually at an NACDL quarterly meeting.

Nomination Guidelines

Nominations may be made by any individual or group and must include:

- the name, title, address and phone number of the nominated person/group
- the name, title, address and phone number of the nominating person/group
- a summary, not to exceed two (2) single-spaced pages, of:
  - the problems that exist(ed) in the relevant indigent defense system
  - the efforts made by the nominee to improve the system (e.g., coalitions formed, legislation proposed, task forces created, litigation initiated)
  - the number of years the nominee has been involved in efforts to improve indigent defense and a brief history of the nominee’s career
  - any changes that have been made in the system as a result of the nominee’s efforts.

Any supplementary materials — such as brochures, reports, or news articles — also may be included. Unlimited letters of support may be submitted. Nominations must be postmarked by December 31, 2008, and mailed to: NACDL Champion of Indigent Defense Award, Attn: Maureen Dimino, 1660 L Street, N.W., 12th Floor, Washington, D.C. 20036.

Eligibility and Selection:

The recipient shall be selected by the Co-chairs of the NACDL Indigent Defense Committee upon the recommendation of the Indigent Defense Award Subcommittee. It is not necessary that the nominee be a lawyer; non-lawyer advocates and reformers will be considered. The Co-chairs of the Indigent Defense Committee and the members of the Indigent Defense Award Subcommittee are not eligible to receive this award but may submit nominations.
standard that when examiners disagree, as they do under the traditional approach, they each can manipulate CMS to show that they are right. On the other hand, CMS can contribute to standardization only if the criterion is inflexible enough to settle disagreements that arise under the traditional approach. Nichols fails to realize that unless CMS is more objective than the traditional “I know it when I see it” approach, there is no justification for using CMS to decide that some, but not other, examiners’ conclusions are right.

Nichols to the contrary, CMS is most favorably viewed as an attempt to use statistical empirical studies to formulate a cut-off point of numbers of consecutive matching striae at which the likelihood that another tool would produce toolmarks that do as good a job at matching the evidence toolmark as the toolmarks produced by the suspect tool is so exceedingly small that, for all practical purposes, the suspect tool can be identified as the unique source of the evidence toolmark. Viewed in this way, CMS is a step in the right direction in that, by contrast to the traditional subjective approach, CMS is at least an attempt to establish statistical empirical foundations for firearms and toolmark identification. It is mistake to suggest, however, as Judge Saris did in Monteiro, that the widespread adoption of CMS would solve the scientific problems with firearms and toolmark identification. To the contrary, CMS is a highly imperfect attempt to establish the requisite statistical empirical foundations.

One of the major problems is that the CMS identification criterion applies to striated, but not impression, toolmarks. This limitation is particularly important because the only types of marks that can be used to determine whether a cartridge case was fired from a particular gun — firing pin impressions and breech face marks — are both impression toolmarks. Nichols’ recent statement that CMS applies both to striated toolmarks and to “toolmarks representative of a striated toolmark (such as striated toolmarks on a breech face impressed onto the head of a cartridge case)” contradicts the dichotomy that he and others have long drawn between striated toolmarks to which CMS applies and impression toolmarks to which CMS does not apply. As firearms and toolmark examiner Kristin Tomasetti states, “CMS … can only be applied to striated toolmarks. It cannot be applied to impression toolmarks, which include granular breech face markings.”

A second problem that even proponents of CMS recognize is that the CMS criterion is intended to be applied to individual, rather than subclass, characteristics of toolmarks. Misidentifications will result if, in applying the criterion, examiners mistake subclass characteristics for individual characteristics. CMS does nothing to remedy the lack of strict rules for distinguishing between subclass and individual characteristics or to decrease the difficulty of making this distinction.

A further, extremely substantial problem is that there are no objective standards for the application of the CMS identification criterion. While numbers of consecutive matching striae must be counted to apply the criterion, “line counting is inherently a subjective process.” “Very often, two independent experts will get different results concerning the total number of striae and the number of matching striae.” The absence of agreement implies that the determination of whether the CMS criterion is met in a particular case is likely to be guided by the individual examiner’s subjective sense of whether evidence and test toolmarks match. This tendency is fostered by the attempts of Nichols and others to defuse opposition by insisting that CMS is not an alternative to the traditional approach, but simply a means by which examiners can describe identifications that they have already reached in their mind’s eye.

In addition, the CMS criterion is not derived from relevant and representative databases. The published statistical empirical studies supporting CMS primarily concern the striae produced by .38 Special Smith & Wesson revolvers on .38 special bullets; bullet striae produced by several other types of guns on several other types of bullets have also been considered. Nonetheless, the criterion is intended to apply to striated toolmarks produced by: (1) all types of firearms on all types of ammunition components, and (2) all tools besides toolmarks.

Moreover, the published studies compare single land impressions on pairs of bullets known to have been fired by different guns to conclude that misidentifications cannot result from the application of CMS to single land impressions. However, in order to avoid false negatives, some examiners advocate declaring a match when the sum of the consecutive matching striae on all land impressions satisfies the CMS criterion. Given the study design, however, the most the studies can show is that false positives will not result if a match is declared when the number of consecutive matching striae on a single land impression meets the CMS criterion. The studies cannot rule out the possibility that misidentifications will result from the application of CMS to the total number of consecutive matching striae on all of a bullet’s land impressions.

### Which Approach Is Better?

Neither the traditional approach nor CMS provides adequate statistical empirical foundations for firearms and toolmark identifications. In sum, adequate statistical empirical foundations have not been developed for firearms and toolmark identification. The subjective approach is tantamount to a denial of the need to do the necessary empirical and mathematical work. CMS is a highly incomplete and flawed approach to the problem.

### Notes


3. See United States v. Diaz, slip op. No. CR 05-00167 WHA, 2007 WL 485967 (N.D. Cal. Feb 12, 2007); Commonwealth of Massachusetts v. Meeks and Warner, Nos. 2002-10961, 2003-10575, 2006 WL 2819423 (Mass. Super. Sep 28, 2006). The author testified for the defense at the admissibility hearings in both cases. Cf. United States v. Williams, 506 F.3d 151, 161-62 (2d Cir. 2007) (holding that the district court’s refusal to grant a hearing before admitting firearms and toolmark identification testimony was not an abuse of discretion, but stating that “[w]e do not wish this opinion to be taken as saying that any proffered ballistic expert should be routinely admitted” and that “expert testimony long assumed reliable before Rule 702 must nonetheless be subject to the careful examination that Daubert and Kumho Tire require”).


5. Id. Judge Rakoff had made similar rulings from the bench in June 2008 in United States v. Damian Brown et al., 05 Cr. 538, and the record of the Daubert hearing in Brown was incorporated by reference in the Daubert hearing in Glynn. See Glynn, 2008 WL 4293317 at *1 & n.1; Transcript of trial in United States v. Brown, 05 Cr. 538 (SDNY June 18, 2008) (“Tr.”) at 1387 (June 16, 2008) and 1476, 1479-81 (June 18, 2008). The author testified for the defense and submitted an affidavit at the Daubert hearing in Brown, and a critique of the author’s work (Ronald Nichols, Defending the Scientific Foundations of the Firearms and Toolmark
Identification Discipline: Responding to Recent Challenges, 52(3) J. FORENSIC SCI. 586, 588 (May 2007), and the author’s published response to the critique (Adina Schwartz, Commentary on Nicholas RG Defending the Scientific Foundations of the Firearms and Toolmark Identification Discipline: Responding to Recent Challenges, 52 J. FORENSIC SCI. 1414-15 (Nov. 2007)), were part of the additional evidence introduced at the Daubert hearing in Glynn. See Glynn at *1 n.2. Readers should note that the rulings in Brown and Glynn were reached even though the Second Circuit had held in Williams, 506 F.3d at 161-62, that it was not an abuse of discretion for a judge to refuse to grant a hearing before admitting firearms and toolmark identification testimony.


7. Id. at 20 (stating that “the proposal for this study explicitly precluded the committee from assessing the admissibility of forensic firearms evidence in court”). See also id. at 81 (stating that “it is not the function of the committee to assess the general validity of firearms identification and toolmark examination nor their admissibility in court proceedings”). Cf. id. at 19 (stating that the committee did not take “a stance . . . for or against the validity of firearms identification generally” and acknowledging that “some may argue that . . . [this] amounts to missing the proverbial elephant standing in the room”).

8. United States v. Monteiro, 407 F. Supp. 2d at 364. See also United States v. Glynn at *3 (citing Green, Monteiro and Diaz for the proposition that “three federal judges have addressed the scientific status vel non of ballistics identification testimony, and all three have concluded that, in one respect or another, it does not have sufficient rigor to be regarded as science”). But see United States v. Williams, 506 F.3d 151, 161 (2d Cir. 2007) (holding that the district court’s refusal to grant a hearing before admitting firearms and toolmark identification testimony was not an abuse of discretion); Decision and Order, United States v. Khalid Barnes, 59 04CR. 186, slip op. at 6 (SDNY April 2, 2008) (“In fact, as the government points out, it appears that no other case in this district or elsewhere has ever found the methodology of traditional pattern matching analysis which was utilized in this case to be unreliable.”).


10. 405 F. Supp. 2d at 108-09.

11. NRC Report, supra n.6, at 72.

12. Id. at 3, 81.


15. NRC, Welcome to the National Research Council, supra.


17. Glynn, 2008 WL 4293317, at *5 (stating that the assumption that each gun produces unique toolmarks “has not been put to the rigorous testing that science demands” and that the reproducibility of the toolmarks produced by individual guns had “never [been] proven to a scientific certainty”).

18. Tr., United States v. Brown at 13 (June 9, 2008). But see United States v. Khalid Barnes, slip op. at 7 (stating that the NRC report “while no doubt useful for the commissioned purpose [of deciding on the viability of a national ballistics imaging database] and not irrelevant to the issue of reliability and admissibility of firearms identification evidence, does not identify any new evidence undermining the core premises upon which ballistics analysis is based, nor does it purport to”); Transcript of hearing in United States v. Ronald English et al., Criminal N. 2007 CF 16618 (Superior Ct. District of Columbia Criminal Division-Felony Branch March 20, 2008) (ruling that the NRC Report did not provide a basis for a Frye hearing on the admissibility of firearms and toolmark identification).

19. 407 F. Supp. 2d at 365. See also id. at 365 (describing studies by firearms and toolmark examiners as "recent scientific studies [that] have demonstrated that the underlying principle that firearms leave unique marks on ammunition has continuing viability"); Meeks, 2006 WL 2819423 at 44-45 (finding that although the prosecution experts’ studies were not blind, they confirmed the premise that each gun leaves individual marks on bullets and cartridge cases); id. 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).

20. Id. at 366.

21. 405 F. Supp. 2d at 110. 22. 2008 WL 4293317 at *5. See also Tr., United States v. Brown at 1477 (June 18, 2008) (stating that “the jury could find, not as science, but as a reasonable assumption, that more individual marks are left.”).


24. Id. at *5-6. See also Tr., United States v. Brown at 1477 (reasoning that serious doubts as to reliability and admissibility are created by the fact that “ballistic matching is inherently subjective in nature” with “no established protocol” for determining whether cartridge cases or bullets match). In both Brown and Glynn, Judge Rakoff’s decision to limit firearms and toolmark examiners to testifying that their conclusions were more likely than not, rather than totally excluding this evidence, was based on the questionable assumption that less rigorous admissibility standards apply to expert testimony that is not scientific than to scientific expert testimony. See, e.g., Glynn at *2 (stating that “the more particular standards for scientific evidence need not be met when the testimony offered does not purport to be scientific” and that the distinction between the applicable admissibility standards “was well-illustrated
in Judge Louis Pollack's well-known [Plaza] decisions regarding fingerprinting evidence). But see Fed. R. Evid. advisory committee's note (2000) ("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 purported to be a scientist.").

25. See, e.g., Jeffrey Scott Doyle, Fundamentals of Firearms ID (2001), available at http://www.firearmsid.com/A_FirearmsID.htm ("Studies have shown that no two firearms, even those of the same make and model, will produce the same unique marks on fired bullets and cartridge cases."); Transcript of Hearing in United States v. Kain, No. 03-573-1 (E.D. Pa. February 2004) at 33, 38-39, 73 (Distinguished Association of Firearms and Toolmark Examiners ("AFTE") Member's testimony that he had never "seen or heard of two different tools creating the same exact tool markings.")


29. Joseph J. Masson, Confidence Level Variations in Firearms Identification Through Computerized Technology, 29 (1) AFTE J. 42 (1997). See also Schwartz, A Challenge, supra n.9, at 11-12 (describing how the above-mentioned article was used in cross-examining Masson in the bolt cutter identification case of United States v. Kain, Crim. No. 03-573-1 (E.D. Pa. 2004)).


31. NRC, supra n.6, at 239.


35. 405 F. Supp. 2d at 108.

36. United States v. Glynn, 2008 WL 4293317 at *5. In the first sentence quoted above, Judge Rakoff failed to realize that gun barrels leave markings on bullets, not casings.


38. Biasotti, A Statistical Study, supra n.34, at 37-38 (1959); Ronald G. Nichols, Firearms and Toolmark Identification Criteria: A Review of the Literature, 42 J. Forensic Sci. 466, 467 (1997). See also Monteiro, 407 F. Supp. 2d at 362 (citing Biasotti's figures for the proposition that "there can be a pattern of matching marks on cartridge cases fired from different guns").


40. 405 F. Supp. 2d at 111.

41. Nichols, Defending the Scientific Foundations, supra n.37, at 587-88.

42. Biasotti & Murdock, Criteria for Identification, supra n.26, at 17.

43. Id.


46. Nichols, Defending the Scientific Foundations, supra n.37, at 587. The judge in Diaz issued his opinion on the Daubert challenge to the admissibility of firearms and toolmark identifications before Nichols effectively retracted his claim about the ease of distinguishing between subclass and individual characteristics. See United States v. Diaz, slip op. No. CR 05-00167 WHA, 2007 WL 485967 at *10 (N.D. Cal. Feb 12, 2007) ("At the hearing the defense emphasized that there appeared to be little literature discussing a technique for identifying subclass characteristics. Nichols explained, however, that it was wrong to assume that subclass characteristics were difficult to account for.").


48. Id.

49. Id. at 228.

50. See Tr., United States v. Brown, at 1390 (June 16, 2008); id. at 1474-76, 1479-80 (June 18, 2008).

51. Id. at 1479 (June 18, 2008).

52. Id.


54. See Bonfanti & DeKinder, supra n.13, at 5 ("A probable solution to the [e] problem [of misidentifications resulting from subclass characteristics on firing pin impressions and breech face marks] lies in a comparison of all the marks present on a cartridge case (breech face impressions, firing pin impression, ejector mark, extractor mark, and marks generally by dynamic processes"); Nichols, Defending the Scientific Foundations, supra n.37, at 590 (stating that "firearms and toolmark examiners are aware that [firing pin impressions] are not wholly reliable for identification to a specific firearm").

55. Biasotti, Murdock & Moran, supra n.13, at 553 n.2.


57. Id.

58. The judge in the case, United States v. Saldana, CR 04-415 (A)-PA (C.D. Cal. 2006), denied the defense request for a Daubert hearing and admitted all the firearms and toolmark identification evidence.

59. Nichols, Defending the Scientific Foundations, supra n.37, at 590.


62. See Champod & Evett, supra n.60 (labeling this value "transparency"); C. Champod, D. Baldwin, F. Taroni, & J.S. Buckleton,

63. Nichols, Defending the Scientific Foundations, supra n.37, at 589.

64. 407 F. Supp. 2d at 369.

65. Id. at 370.

66. 2008 WL 4293317 at *4. In Brown, when the prosecution expert referred to the AFTE Theory, Judge Rakoff interjected, “[Y]ou said the definition of sufficient agreement of characteristics is that there are enough – I think ‘enough’ and ‘sufficient’ may be synonyms, but I come back to the question I had before: Is what you are really saying here that if you are in this field long enough, you know it when you see it?” Tr., United States v. Brown, at 1115-16 (June 13, 2008). See also id. at 1477 (June 18, 2008) (statement by Judge Rakoff that “even the people most favorable to ballistics evidence say that there is no established protocol for [determining whether test and evidence toolmarks match]. It has to be ‘sufficient,’ a term that is inherently vague.”).

67. In United States v. Glynn, Judge Rakoff referred to CMS and stated that, “Although attempts have been made to introduce … minimum standards and ‘protocols’ into ballistics analysis, such attempts have not yet met with general acceptance and, in any event, were not applied by the examiners in Brown and Glynn.” Glynn, supra at *5 (footnote omitted).

68. Biasotti, A Statistical Study, supra n.34, at 34.

69. Biasotti, Principles of Evidence Evaluation, supra n.61, at 429. See United States v. Glynn, 2008 WL 4293317 at *4 & n.8 (citing Biasotti’s statement and the NRC Report as support for the proposition that “the literature confirms [that] ballistics opinions are significantly subjective”).


71. Nichols, Defending the Scientific Foundations, supra n.37, at 590.

72. Id.


74. Nichols, Defending the Scientific Foundations, supra n.37, at 590.


78. Nichols, Defending the Scientific Foundations, supra n.37, at 591.


80. Tomasetti, supra n.79, at 298.


82. Tomasetti, supra n.79, at 298. See also Charles Meyers, Some Basic Bullet Striae Considerations, 34(2) AFTE J. 158, 158-59 (Spring 2002); Bunch, supra n.70, at 959.

83. Werner Deinet, Comments on the Application of Theoretical Probability Models Including Bayes Theorem in Forensic Science Relating Firearms and Toolmarks, 39(1) AFTE J. 4, 6 (Winter 2007).


85. See, e.g., Nichols, Defending the Scientific Foundations, supra n.37, at 590; Bruce Moran, Comments and Clarification of Responses From a Member of the AFTE 2001 Criteria for Identification of Toolmarks Discussion Panel, 35 (1) AFTE J. 55, 61 (Winter 2003); Moran, A Report, supra n.47, at 231; Biasotti, Murdock & Moran, supra n.13 (stating, in an edition written after Biasotti’s death: “It should, therefore, be clear that the application of the quantitative CMS criteria is not a different method than pattern matching, but is merely a quantitative way to describe the extent of striated pattern matching agreement.”).

86. See Walsh & Wevers, supra n.75, at 4.


88. Cf. Champod, Baldwin, Taroni, & Buckleton, supra n.62, at 313 (“What we really needed [in order to convert CMS into a rigorous probability model] was the highest number of CMS noted in the whole bullet. We cannot see how to get this data from Mr. Biasotti’s paper. …”).

© Adina Schwartz, 2008. All rights reserved.

Inside NACDL

Continued from page 7

operations on relations with America’s Muslim community and gauge the effect of these tactics upon public confidence in law enforcement; (3) expose the practical impact of massive electronic surveillance; and (4) identify pressing national security issues that must be addressed by Congress, the courts, and the legal profession.

NACDL recognizes that as the nation confronts the shadowy threat of terrorism, a threat that can manifest itself whenever a solitary, suicidal lunatic straps a bomb to his back, the national inclination to suppress constitutional liberty remains an irresistible lure for politicians of all stripes. American lawyers, especially the criminal defense bar, have a special duty to preserve and protect fundamental liberties. Roberto Bolano, the Chilean author of the novel By Night in Chile, described how many in the Pinochet era accommodated themselves to brutality: “The answer was simple. Because with time, vigilance tends to relax, because all horrors are dulled by routine.”

NACDL pledges not to let this happen in America. Liberty’s Last Champion will never accept extra-constitutional measures as routine.

About the Author

Adina Schwartz

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

Inside NACDL

Continued from page 7