UNITED STATES DISTRICT COURT
FOR THE DISTRICT OF MASSACHUSETTS

UNITED STATES OF AMERICA,	)	
	)	
v.	)	**CRIMINAL. NO. 02-10301-NG**
	)	
DARRYL GREEN, et al.,	)	
Defendants	)	

GERTNER, D.J.:

TABLE OF CONTENTS

MEMORANDUM AND ORDER RE:
MOTION TO EXCLUDE BALLISTICS TESTIMONY
December 20, 2005

I. INTRODUCTION . . . . . -1-

II. FACTS . . . . . -7-
- A. Examiner Had No Certification by Professional Organization . . . . . -7-
- B. Premise of the Field: That the Markings on Each Firearm Are Unique . . . . . -9-
- C. The Examination: Sorting out Class, Subclass, Individual, and Accidental Characteristics with No National Standards . . . . . -11-
- D. The Examination: No Notes, No Drawings, No Photographs . . . . . -16-
- E. The Examination: Observer Bias and Lack of Blind Testing . . . . . -21-
- F. The Examiner: No Data on Error Rates . . . . . -21-
- G. Computerized Database: The IBIS System . . . . . -23-
- H. Testimony of Defense Expert David Lamagna . . . . -23-

III. LEGAL FRAMEWORK . . . . . -25-

IV. THE CASE AT BAR . . . . . -30-
- A. Reliability and Testing . . . . . -30-
- B. Error Rates and Proficiency Testing . . . . . -34-
- C. Significance of Defense Expert and Availability of Cross-Examination . . . . . -35-
- D. Significance of Precedent . . . . . -36-

V. CONCLUSION . . . . . -39-

AUTHENTICATED U.S. GOVERNMENT INFORMATION GPO

UNITED STATES DISTRICT COURT
FOR THE DISTRICT OF MASSACHUSETTS

UNITED STATES OF AMERICA,	)	
	)	
v.	)	**CRIMINAL. NO. 02-10301-NG**
	)	
DARRYL GREEN, et al.,	)	
Defendants	)	

GERTNER, D.J.:

MEMORANDUM AND ORDER RE:
MOTION TO EXCLUDE BALLISTICS TESTIMONY
December 20, 2005

I. INTRODUCTION

Defendants Jonathan Hart and Edward Washington challenge the admissibility of forensic ballistics identification evidence pursuant to Fed. R. Evid. 702. Significantly, they do so in the context of a case involving extremely serious allegations, namely racketeering, assault in aid of racketeering (pursuant to 18 U.S.C. § 1962(c), § 1962(d), 18 U.S.C. § 1959(a)(3)) and various gun charges (pursuant to 18 U.S.C. § 924(c)).[1]

On September 9, 2000, Boston police detectives collected eight spent .380 caliber shell casings on the sidewalk opposite 249 Harvard Street in Boston. On September 16, 2000, following the shooting of Richard Green, detectives collected six spent .380 caliber shell casings at 870 Blue Hill Avenue in Boston. More than a year later, on September 20, 2001, detectives found a

[1] While the trials of Hart and Washington have been severed from the remaining defendants in the case at bar, Branden Morris, Darryl Green, and Torrance Green, I can only assume that similar issues will be raised in those cases. If convicted, Morris and Green face the death penalty.

loaded Hi Point, .380 caliber pistol [“Hi Point pistol”] in the front yard at 6 Esmond Street in Boston.

The government seeks to introduce ballistics testimony from Sergeant Detective James O’Shea [“O’Shea”] of the Boston Police Department. O’Shea examined the evidence and concluded that all of the shell casings came from the same weapon, and further, that the weapon was the Hi Point pistol found in front of 6 Esmond Street. Indeed, O’Shea declared that this match could be made “to the exclusion of every other firearm in the world.” (Daubert Hr’g Tr. 20, 60, Oct. 5, 2005.)

That conclusion, needless to say, is extraordinary, particularly given O'Shea's data and methods. After hearings were held over several days, at which I received testimony from O’Shea and a defense ballistics witness, David Lamagna, I found the following (as described more extensively below):

Although O’Shea has seven years of experience in the Boston Police Ballistics unit, neither he nor the laboratory in which he worked has been certified by any professional organization. He has worked on hundreds of cases, but has never been formally tested by a neutral proficiency examiner. Nor could he cite any reliable report describing his error rates, that of his laboratory, or indeed, that of the field.

The shell casings found at the two sites did not exactly match the shell casings test-fired from the Hi Point gun found on

Esmond Street. In firearm toolmark comparisons, exact matches are rare. The examiner has to exercise his judgment as to which marks are unique to the weapon in question, and which are not.

In fact, shell casings have myriad markings, some of which appear on all casings from the same type of weapon ("class characteristics") or those manufactured at the same time ("sub-class characteristics"). Others are arguably unique to a given weapon ("individual characteristics") or are unique to a single firing ("accidental characteristics"). The task of telling them apart is not an easy one: Even if the marks on all of the casings are the *same*, this does not necessarily mean they came from the *same gun*. Similar marks could reflect class or sub-class characteristics, which would define large numbers of guns manufactured by a given company. Just because the marks on the casings are *different* does not mean that they came 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 distinguishing class and sub-class characteristics from individual ones, O'Shea did not have many resources to rely on. He conceded, over and over again, that he relied mainly on his subjective judgment. There were no reference materials of any specificity, no national or even local database on which he relied. And although he relied on his past experience with these

weapons, he had no notes or pictures memorializing his past observations. He could have contacted the Hi Point manufacturer directly to ask about how the particular gun he was examining was manufactured or obtain diagrams or photographs of its features, but he did not.

The only weapon he 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. He 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."

O'Shea reviewed the evidence when it was submitted to him five years ago, but took no notes, recorded no measurements, made no photographs, and drew no diagrams. He agreed that to the extent there were protocols for toolmark examination, he did not follow them in this case. He did not take photographs of the evidence until a week before his testimony at the Daubert hearing.

Defendants offered the testimony of David Lamagna, which had its own substantial weaknesses. Lamagna's qualifications were

strongly challenged by the government, but his testimony corroborated many of the deficiencies that had already been established during the cross-examination of O'Shea. Lamagna had no laboratory, and although he had an advanced degree in material science, his ballistics experience derived largely from armorer courses offered by gun manufacturers.

In any case, notwithstanding all of the serious deficiencies, the problem for the defense is that every single court post-Daubert has admitted this testimony, sometimes without any searching review, much less a hearing.[2] In addition, whatever Lamagna's deficiencies as a witness, it is somewhat significant that he did have access to the underlying physical evidence. Since O'Shea's examination did not involve destructive testing, the defendants could offer alternative results to the jury. O'Shea's expertise could arguably be challenged in a way that would be accessible to the jury both on cross-examination and via defense testimony.

Given this precedent, and notwithstanding my serious reservations, I feel compelled to allow O'Shea to testify about

[2] In strong contrast, in United States v. Monteiro, No. 03-10329 (D. Mass. Nov. 28, 2005). after a six day hearing, Judge Saris in the District of Massachusetts recently concluded that Sergeant Douglas Weddleton of the Massachusetts State Police, a toolmark examiner, "did not follow the established standards in the toolmark identification field with respect to documentation and peer review of his results." Id. at 2. She ordered that the government comply with these standards and provide the defense with the necessary documentation and peer review, as a prerequisite to his testimony. Judge Saris reserved the ultimate question with which this decision is concerned, the "reliability of the standard methodology in the field," the kinds of conclusions this expert and like experts may interpose. Id. at 16.

his observations of the shell casings found at the site of the September 9 shooting at 249 Harvard Street and the September 16 Richard Green shooting on Blue Hill Avenue, and about his comparison of those casings to the suspect Hi Point weapon. However, I will also follow the procedure I used in United States v. Hines, 55 F. Supp. 2d 62 (D. Mass. 1999), with regard to handwriting analysis.[3] As in Hines, O'Shea may only describe and explain the ways in which the earlier casings are similar to the shell casings test-fired from the Hi Point pistol found a year later. I will not allow him to conclude that the shell casings come from a specific Hi Point pistol "to the exclusion of every other firearm in the world."[4] That conclusion -- that there is a definitive match -- stretches well beyond O'Shea's data and methodology.

I reluctantly come to the above conclusion because of my confidence that any other decision will be rejected by appellate courts, in light of precedents across the country, regardless of the findings I have made. While I recognize that the Daubert-Kumho standard does not require the illusory perfection of a

---

[3] In Hines, this Court concluded that although a handwriting expert could legitimately testify to similarities and dissimilarities between handwriting samples, the expert could not testify to an exact match to the exclusion of all other samples. As with the toolmark analysis in this case, the Court found in Hines that handwriting evidence was not sufficiently reliable to justify an expert's ultimate conclusion announcing a match.

[4] Compare DNA testing: DNA evidence of a "match" is only admitted with statistical evidence of the probability of a coincidental match, not as a definitive statement.

television show (CSI, this wasn't),[5] 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.[6]

II. FACTS

A. Examiner Had No Certification by Professional Organization

The government's proffered expert, Sergeant Detective O'Shea, has worked in the Boston Police ballistics unit for seven years (since 1998). He received "armorer's training" in a Smith & Wesson revolver and two types of Glock pistols. He served as an apprentice to another detective for six to twelve months, although the qualifications of the other detective were not described to the Court. O'Shea's training consisted of observing his mentor's work and "appl[ying] . . . best practices from the

[5] CSI ("Crime Scene Investigation") is a dramatic television series on CBS about a team of high-tech forensic specialists who use their technological and forensic expertise to solve crimes.

[6] Indeed, recent reexaminations of relatively established forensic testimony have produced striking results. Saks and Koehler, for example, report that forensic testing errors were responsible for wrongful convictions in 63% of the 86 DNA Exoneration cases reported by the Innocence Project at Cardozo Law School. Michael Saks and Jonathan Koehler, The Coming Paradigm Shift in Forensic Identification Science, 309 Science 892 (2005). This only reinforces the importance of careful analysis of expert testimony in this case.

lab to my work." (Daubert Hr'g Tr. 5-7, Oct. 5, 2005.) O'Shea estimated that he has done "hundreds" of ballistics examinations in his career and testified in a similar number of cases. Id. at 7. However, as described below, the record does not indicate how many of these "hundreds" of examinations were accurate.

O'Shea has never received proficiency testing from any neutral entity. While the Association of Firearm and Toolmark Examiners (AFTE) certifies ballistics examiners and O'Shea has attended the annual training seminar for the New England AFTE chapter every year except this past year (Daubert Hr'g Tr. 36, Oct. 5, 2005), he has never been AFTE-certified. (Daubert Hr'g Tr. 82, Nov. 2, 2005.) He is not even an AFTE member, nor has he bothered to take their exam.[7] Id. And whatever protocols the AFTE has established for examining ballistics, as Judge Saris described in United States v. Monteiro, O'Shea did not follow them at the time of his initial examination. Nor is O'Shea's laboratory, the Boston Police Laboratory, certified by any organization.[8]

---

[7] The AFTE publishes a journal that is peer-reviewed by other members of the field, but the "field" consists entirely of individuals who work for law enforcement agencies. (Daubert Hr'g Tr. 41, Oct. 5, 2005.) In contrast, the DNA-typing "field" involves neutral academics as well as law enforcement personnel.

[8] The laboratory is in the process of obtaining certification from the American Society of Crime Lab Directors, ASCLAD, but this process had not even begun at the time of the initial examination in this case. (Daubert Hr'g Tr. 84, Nov. 2, 2005.)

B. Premise of the Field: That the Markings on Each Firearm Are Unique

The premise of the field -- which the defendants and others have contested -- is that the surface contours of each firearm are unique, even in an age of mechanized production. (Daubert Hr'g Tr. 29, Nov. 2, 2005.) As guns are produced, the "tool is dulled" because "it's metal cutting metal," and the result is that "the edge won't be as deep when it cuts the second barrel. It will be a little wider." (Daubert Hr'g Tr. 8, Oct. 5, 2005.) The marks are enhanced when the breech face[9] of the gun is roughened with sandpaper; because "sand falls off in that process," O'Shea concluded that each face will have a different surface. Id. at 8-13.

To be sure, this description, while perhaps accurate in the past, may not describe modern-day gun manufacturing processes. In any case, the government offered evidence that the Hi Point pistol in question, manufactured in 1998, was, indeed, hand-sanded.[10]

---

[9] The breech face is the inside rear of a gun, where the bullet rests prior to being fired.

[10] The government offered an affidavit from Tom Deeb, President and CEO of Hi Point Firearms, dated October 3, 2005. The government provided him with the serial number for the Hi Point recovered on September 20, 2001. Based on his records, he was able to determine the manufacturing process that had been used on the weapon. Significantly, he noted that the breech face of the firearm was finished by "hand sanding with an 80 grit sanding belt." (I admitted the affidavit with the last two sentences excluded. The last two sentences were challenged by the defense –- that each firearm is uniquely marked, that those marks are transferred to the cartridge cases, and that a qualified examiner can determine whether a particular shell casing was fired from a particular weapon.)

When a gun is fired, some of the gun's features are transferred to the shell casings, creating patterns of striae (scratch marks) as the cartridge casing leaves the gun. The gases produced when the primer explodes and gunpowder burns cause the casing to expand in all directions; as a result, markings from the breech face of the gun are imprinted to some degree onto the casing. (Daubert Hr'g Tr. 15, Nov. 2, 2005.) In addition, in the process of igniting, the firing pin creates an imprint on the cartridge case. On semiautomatic firearms, a metal spring called the "extractor" can also leave individualized markings. Id. at 18. There are also chamber marks on fired casings, left by the sides of the firing chamber as the casing slides through the gun. Id. at 32-33.

There is no question that there are many marks on shell casings, from all of these sources -- production process, firing pin, breech face, etc. But 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.

O'Shea did not consult with Deeb at the time he initially examined the Hi Point.

C. The Examination: Sorting out Class, Subclass, Individual, and Accidental Characteristics with No National Standards

The goal of the ballistics examination is to distinguish between class and sub-class characteristics on the one hand, which include large numbers of guns, and individual characteristics on the other, which ostensibly apply to the particular gun in question. Class characteristics are "design features, they're what the manufacturer intends before the first piece of metal is processed." (Daubert Hr'g Tr. 16, Oct. 5, 2005.) Presumably, class characteristics narrow the identification to a given manufacturer, but not to an individual gun. Thousands of firearms may share the same class characteristics. (Daubert Hr'g Tr. 16-17, Nov. 2, 2005.) Sub-class characteristics are markings that temporarily become part of the manufacturing process and therefore create a marking on perhaps hundreds of weapons in a given production run, though they are not a permanent feature of the design. (Daubert Hr'g Tr. 23-24, Oct. 5, 2005.) In effect, sub-class characteristics indicate an imperfection in the method used to produce a limited number of firearms. (Daubert Hr'g Tr. 22-23, Nov. 2, 2005.)

In contrast, the individual characteristics are "imparted from the actual piece, the actual tool" during production. (Daubert Hr'g Tr. 15-16, Oct. 5, 2005.) Finally, there are

accidental characteristics that may be found in shell casings fired from the same gun.[11]

The examiner's task is further complicated by the fact that an individual gun's markings change over time; marks present at one period may not be there at another (a fact that may be relevant in this case, where O'Shea seeks to compare shell casings fired at one point with casings test-fired from a gun found a year later).[12] Moreover, two-dimensional analysis may be misleading. Although the pattern may look like a class characteristic, the particular depth of the line examined could be an individual one. Plainly, confusing individual characteristics with class or sub-class ones could lead to false negatives, as well as false positives.

---

[11] Two shells may be different, even if from the same gun. Successive firings, however, can produce different marks. The firing pin may not retract fast enough, "turn[ing] that round impression into a T-hole shape." (Daubert Hr'g Tr. 18, Oct. 5, 2005.) For example, on one shell, the firing pin may have made a deeper impression on one occasion, resulting in more circles on that shell than on another shell fired from the same gun. (Daubert Hr'g Tr. 45, Nov. 2, 2005.)

[12] In addition, O'Shea acknowledged a study published in the AFTE journal by the Alabama Department of Forensic Sciences finding that even with higher quality guns (with harder metals than the Hi Point), the metal wore down over time such that cartridges from earlier firings could not be matched with cartridges from later firings. (Daubert Hr'g Tr. 63, Oct. 5, 2005.) (This article is Exhibit 9.) However, O'Shea said that this type of wear over time was not consistent with his personal experience. Id. at 64. According to O'Shea, a test done by the Boston Police Department of 2200 Glock guns found that the markings did not change notably over time. In contrast, in a study of Glock firearms by the Georgia Bureau of Investigation, examiners fired a number of shells from each firearm and had difficulty matching the projectiles with the guns. Id. at 48. In any case, O'Shea acknowledged that the Hi Point, one of the least expensive guns there is, wears differently than a Glock.

Finally, the analysis is affected by the kind of gun in question. A Hi Point is one of the cheapest guns made. It is not a particularly high quality gun, in part because it is made with softer steel (10/10 steel) than other firearms. Because it is soft, it is more susceptible to wear over time than the steel in other firearms. (Daubert Hr'g Tr. 75, Oct. 5, 2005.)

In distinguishing class and sub-class characteristics from individual ones, O'Shea had little upon which to rely. There are reference works that discuss some of the known class and sub-class characteristics but "most of the time" an examiner would be deciding whether a mark was an individual versus sub-class or class characteristic on his own.[13] (Daubert Hr'g Tr. 23-25, Nov. 2, 2005.) The FBI publishes "general rifling characteristics" each year which help identify class characteristics to a limited degree (e.g. relative position of firing pin and extractor), although no such report was introduced in the instant case. (Daubert Hr'g Tr. 56, Oct. 5, 2005.) While O'Shea said that it helps to "know some of the manufacturing process" and it "never hurts to pick up the telephone" to call the manufacturer for this information (Daubert Hr'g Tr. 24, Oct. 5, 2005), there is no evidence that he did so in this case.

Nor did O'Shea do anything to systematize his own past experience. He never kept any written record of the

[13] Individual marks may change over time. (Daubert Hr'g Tr. 26, Nov. 2, 2005.) There are a "mountain" of reference works, O'Shea noted, although none were offered. In any case, he conceded over and over again that in the end, whether a mark is an individual mark or a sub-class mark is a subjective determination. Id. at 28.

characteristics of the guns he has examined, although he reported that he has conducted "hundreds" of forensic exams since 1998.[14] In effect, to decide if something could be a sub-class or class characteristic, he just compares the image in front of him to what he *remembers* from all those previous exams.[15] (Daubert Hr'g Tr. 38-40, Nov. 2, 2005.)

---

[14] This methodology is of particular concern in distinguishing sub-class from individual characteristics. The first time an examiner observes a particular sub-class characteristic, he may assume it is an individual characteristic.

[15]

Q. Now, let me show you also Exhibit 2 [with the casings found at the Green shooting and at 249 Harvard Street], and if I got this right, on the right side you see what appears to be an upside down checkmark type . . . . On the left side, do you see what appears to be a checkmark type impression?
A. I see part of it appears to be an impression, part of it appears to be a scratch or a striae.
Q. Right. It's a mark on the shell casing that's produced by the firing pin, correct?
A. I don't know that it is . . . .
Q. Well, what is it?
A. I don't know that it is produced by the firing pin.
Q. Well, what significance, do you give it any significance?
A. Do I, no.
Q. Why?
A. It appears to be a mark that's accidental in nature, whether it was by the manufacturer, the primer, or it was scratched prior to being placed in there.
Q. How do you know that?
A. I don't know that, that's why I say or. . . .
Q. What basis, what studies, what database, what anything do you have that would allow you to say I'm not going to count that, I'm still going to call it a match even though the one on the right does not have that mark?
A. None.
Q. So it's just your opinion? You determine which marks you're going to pay attention to and which ones you're not, correct?
A. Correct.

(Daubert Hr'g Tr. 92-93, Oct. 5, 2005.)

O'Shea did not examine any .380 Hi Point pistols other than the suspect pistol in this case. He had shell casings from test-firings from four other Hi Point firearms, which happened to be in Boston police custody (Daubert Hr'g Tr. 66, Nov. 2, 2005), but he knew which casings had come from the suspect weapon, and which had come from other guns.[16] He examined the other shell casings just to provide some context, to help him identify class characteristics of the Hi Point. The problems is that O'Shea did not look up the serial numbers of the guns or otherwise determine whether those sample guns were manufactured near the time of the suspect gun, a factor potentially relevant to noting sub-class characteristics. (See Daubert Hr'g Tr. 70, Nov. 2, 2005.) In fact, when he examined the casings from the other three guns already in police custody, the only class characteristic he could discern was the shape of the firing pin, and "probably the horizontal or parallel breech face marks." Id. But even the "firing pin" observation was equivocal; a firing pin may strike a shell differently each time.[17] Id. at 74. In any event, O'Shea

---

[16]

Q. Do you have any kind of database that would allow you, for instance, to go in and see if Lorsen [another gun manufacturer] has weapons in which the top of the firing pin is not quite round?

A. No.

Q. Do you have any kind of database that would allow you to look and see how many Hi Points that may have a top that's not quite round?

A. No.

(Daubert Hr'g Tr. 91, Oct. 5, 2005.)

[17]

Q. Okay, so, what did you consider to be sub-class characteristics with respect to firing pin impressions of each of these

acknowledged that the test-fired casings from the suspect weapon had many fewer marks on them than those test-fired from the other weapons already in police custody. Id. at 72.

D. The Examination: No Notes, No Drawings, No Photographs

O'Shea used a comparison microscope to analyze the evidence, which allowed him to see two cartridges through a single eyepiece. It was a 10 power microscope with base optics of 8 to 12.5, so the magnification was 80-125 times. Adjustable lighting to the side of the scope illuminated the samples.[18] (Daubert Hr'g Tr. 13-15, Oct. 5, 2005.)

O'Shea looked for a readily identifiable feature on one cartridge and lined up the second cartridge to sit in the same position, based on that same feature. O'Shea took no pictures and made no notes of his work when he examined the evidence in 2001. Id. at 45. He measured some of the striae with a reticle on the microscope, but did not record the measurements. Id. at 60-61. In any event, with the reticle, he could only measure

---

two test-fires [one of the casings found in 2000 and one of the casings from guns in BPD position]?
A. Off the top of my head, I didn't see anything.
Q. What were the individual characteristics?
A. The fact that this interior mark on the left is no longer round, I consider that to be an individual characteristic.

(Daubert Hr'g Tr. 74, Nov. 2, 2005.)

[18] The comparison microscope is a relatively crude tool. O'Shea acknowledged that the Boston Police Department does not use a scanning electronic front microscope, laser microscope, or white light microscope. He said this was because of the "cost, most likely," but also that he was not sure how repeatable or reliable that technology was. (Daubert Hr'g Tr. 38, Oct. 5, 2005.) O'Shea was not aware of any police or law enforcement lab, other than IBIS (discussed below) that used that technology.

length and width, not three-dimensional contour. (Daubert Hr'g Tr. 75, Nov. 2, 2005.) It was not until the week of the hearing before this Court that O'Shea finally took pictures of some of the evidence, five years after his initial exam, to prepare for his testimony in the Daubert hearing.

O'Shea acknowledged the limitations of the photographs: The lighting is not as good on the photos as it was when he examined the evidence. Moreover, there is a dividing line on the microscope that can be manipulated to aid comparison; a still photo does not permit that manipulation. The Court was unable to see the matching marks that O'Shea pointed out.

In any event, when O'Shea lined up the casings that were found with those test-fired from the Hi Point recovered on Esmond Street, he conceded there was no exact match, and that he had to exercise judgment as to which marks were significant and which to discount or ignore.

Asked about the standards for determining a match that would guide his judgment, O'Shea's testimony was either tautological or wholly subjective. The tautological: He said "[t]he standard that the identifiable features, the repeatable features, that are observed under the microscope, obviously have to be such that it's identified to one firearm only to the exclusion of all others. . . ." (Daubert Hr'g Tr. 20, Oct. 5, 2005.) The subjective: ". . . it has to present with individual characteristics that satisfy me in the end that it couldn't have

come from any other firearm." Id. Indeed, he repeated this point over and over again.[19]

In effect, there are no national standards to be applied to evaluate how many marks must match.[20] (Daubert Hr'g Tr. 55, Oct. 5, 2005; Tr. 77, Nov. 2, 2005.) O'Shea cited the AFTE's published "Theory of Identification as it Relates to Toolmarks" (Ex. 1), which requires a "sufficient agreement" between samples in order to declare a match. O'Shea said there were others he could consult with to define "sufficient agreement," but "in the end, it's my judgment." (Daubert Hr'g Tr. 42, Oct. 5, 2005.) He did not describe what "sufficient agreement" was, though he did state that it "basically comes down to counting lines." (Daubert Hr'g Tr. 30, Nov. 2, 2005.) In the Boston Police Department, there is a "suggestion" about the proper procedure in the manual,

---

[19] Daubert Hr'g Tr. 22, 55, 86, Oct. 5, 2005. When the Court asked how O'Shea could know by looking at an individual firearm whether a mark was an individual or sub-class characteristic without examining hundreds of firearms, his response was less than clear: "It's absolutely correct, and that's why I would never make an identification based on that type of mark. It requires some caution to discount." Id. at 24. Nor is it clear how he could ever distinguish individual from class or sub-class characteristics without calling the manufacturer about each individual marking. He acknowledged that the coarser marks on a breech face could be either class or individual characteristics. His determination of a match was "a subjective judgment, where they appeared and how they appeared." (Daubert Hr'g Tr. 37, Nov. 2, 2005.) And even more vaguely: It is not individual marks that O'Shea finds significant, it's "the totality . . . The relationship of them one to another" that is significant. Id. at 59.

[20] Indeed, O'Shea acknowledged the difficulty of applying the subjective standards. For example, he testified that the Boston Police Department conducted a test-firing of all of its Glock pistols and found that Glocks produce a "readily identifiable firing pin impression." (Daubert Hr'g Tr. 23, Oct. 5, 2005.) However, this impression is a *class* characteristic, rather than an individual one, so "it became apparent to me that there was a danger in identifying Glock pistol to Glock pistol to Glock pistol just because of that class characteristic, so having become aware of it at an early stage, I'm very, very cautious about what I'll identify as an individual characteristic and from there a match." Id. There is no indication that a similar test was performed on Hi Points, and it could well be that Hi Points pose similar challenges.

but it is not incumbent on the examiners to follow it. "[E]ach examination is going to be different than the one before it." (Daubert Hr'g Tr. 83-84, Oct. 5, 2005.) O'Shea also mentioned the Consecutive Matching Striae ("CMS") criteria for declaring a match. Although he said he "believed" he used CMS in this case, he could only vaguely describe what the criteria were and how they were applied here. He could not even state clearly how many similar lines were required for a match.[21] (Daubert Hr'g Tr. 31-32, Nov. 2, 2005.) He acknowledged that, like the "sufficient agreement" standard, the CMS criterion are subjective. Id. at 31.

Specifically, looking at the photo marked Exhibit 6,[22] O'Shea acknowledged that the circles on the two side-by-side samples were not identical. Id. at 87. He said he could match them based on the shape at the tip of the firing pin, which was a polygon, even though the machining process would have made it round. Id. at 90. However, he acknowledged that he did not have a database from which he could determine whether other

---

[21]

Q. Does the Boston Police Department adhere to CMS identification criteria?
A. I believe we do.
Q. Okay. Can you tell us what those -- what CMS identification criteria are?
A. It's, you're making me guess now, but if there's one group of matching striae, I believe it has to be six consecutive matching striae. If there's two, then you can drop down a little bit in the number.

(Daubert Hr'g Tr. 31, Nov. 2, 2005.)

[22] Exhibit 6 depicts shell casings from the September 16, 2000 shooting and a test-fire.

inexpensive firearms, such as those manufactured by Lorsen, also produced a firing pin that was not quite round. Id. at 91. Nor did he have the ability in his lab to measure the exact shape of the tip so as to compare it to others.

Looking at the photo marked Exhibit 2,[23] O'Shea acknowledged a mark on one sample that looked like a checkmark, which was not present on the other sample. He said that he did not know whether it was produced by the firing pin and that he did not attach any significance to it ("[i]t appears to be a mark that's accidental in nature"), though he did not know its exact source. When asked "What basis, what studies, what database, what anything do you have that would allow you to say I'm not going to count that, I'm still going to call it a match even though the one on the right does not have that mark?" O'Shea responded "None"; it was just his opinion. Id. at 92-93.

The instant examination was complicated by the nature of the gun, a Hi Point. The breech faces on the Hi Point .380's O'Shea examined were relatively smooth; there were not very many marks. (Daubert Hr'g Tr. 57, Nov. 2, 2005). O'Shea conceded that there are not as many toolmarks on the shell casings in this case as there sometimes are on other casings. Id. at 71-72. The fewer the markings, the less data the toolmark examiner has to determine a match. Again, other gun manufacturers -- particularly of inexpensive guns like the Hi Point -- produce

[23] Exhibit 2 displays casings from the September 9 and September 16, 2000 shootings.

weapons with similar class characteristics, but O'Shea did not compare the casings in this case to those from other inexpensive guns.

E. The Examination: Observer Bias and Lack of Blind Testing

While O'Shea indicated that he did not know the source of the samples or firearms at the time he did his examination, he was not presented with more than one firearm. (Daubert Hr'g Tr. 85-86, Oct. 5, 2005.) He had test-fires from three guns in Boston Police custody, but, as described above, he used these test-fires to provide context -- a sample of other Hi Point casings -- not as alternative matches. In effect, the examination was an evidence show-up (do these casings come from this gun?), not an evidence line-up (from which gun do these casings come?).

F. The Examiner: No Data on Error Rates

If the basis for O'Shea's subjective observations will not be apparent to a factfinder, at the very least, the factfinder should have some data to determine how much confidence to place in O'Shea's powers of observation -- proficiency testing, error rates, or certification. Asked about error rates, O'Shea testified that for false positives, "[t]he stated error rate for false identifications I believe is greater than one percent, less than two. It's probably around one and a half percent." Id. at

40.[24] But there is no indication of the source of this statistic. Indeed, O'Shea conceded that there have been no controlled studies to evaluate the error rate of the field.

Nor was O'Shea himself certified as an examiner by any outside organization. He has never taken a certification exam. Moreover, on one occasion, he informally filled out proficiency tests from the Comprehensive Testing Service (a firm in Florida), but again, he never submitted the test to the Service to be graded. Another examiner from the Boston Police Department did submit his results, which O'Shea reported were correct. O'Shea claimed that his results happened to correspond to that of the other examiner. There is no corroboration of the evidence.

Nor is it clear that O'Shea's examination was reviewed by another, unbiased examiner. Under Boston Police policy, more than one person must examine the evidence to make an identification -- that is, someone must review the initial identification. (Daubert Hr'g Tr. 50, Nov. 2, 2005.) There are "quite often" disagreements. If they cannot reconcile their differences, there will not be an identification. Id. at 50-51. But there are no notes of this subsequent review and no indication it was blind testing. The second examiner knew that O'Shea had already identified a match.

[24] Q: So what statistics are there that determine the likelihood that the characteristics would match a firearm of a certain type? . . . A. None that I'm aware of. (Daubert Hr'g Tr. 75, Nov. 2, 2005.)

G. Computerized Database: The IBIS System

O'Shea also used the Integratable Ballistic Identification System (IBIS) in his comparison, although the government represented that it would not offer IBIS results. A national computer database, IBIS allows examiners to identify the most likely matches for the evidence in a given case. IBIS uses a laser measuring device to evaluate shell casings and provides the examiner with a list of possible matches. (Daubert Hr'g Tr. 39, Oct. 5, 2005.)

In fact, the IBIS system has been widely criticized. Its efficacy is limited by the detail with which police departments have scanned old shell casings into the computer and the accuracy of the mathematical algorithms used to compare casings. Id. at 68. As with the individual examinations, no evidence was presented about the accuracy of the IBIS matches.

In any event, O'Shea acknowledged that even if the computer suggests numerous possible matches, he will not bother to check them all. Id. at 73. That is, once he decides he has found a match, he will not eliminate all other alternatives by exhausting the IBIS-generated list of potential matches.

H. Testimony of Defense Expert David Lamagna

Defense expert David Lamagna provided limited insight into the problems with O'Shea's methodology. To say the least, Lamagna's credentials left something to be desired. He proclaimed expertise in dozens of fields, but his toolmark

experience derived largely from short courses offered by gun manufacturers. (Daubert Hr'g Tr. 85-90, 7-8; Nov. 21, 2005.) He is not professionally affiliated with any forensic organization and has only published one article, and even that did not appear in a peer-reviewed journal. Id. at 90-91.

Nevertheless, I credit his testimony only to the extent it confirmed O'Shea's description of several features of the field which confound the identification. To note a few: first, wear on a firearm over time may affect the marks it leaves. Id. at 27-28. This problem is exacerbated by the soft, 10/10 steel from which Hi Points are crafted. Id. at 36-38. Second, marks on two shell casings from the same gun may vary because not all marks will necessarily imprint onto the casing every time the gun is fired. Id. at 29. Third, O'Shea did not engage in three-dimensional analysis of the toolmarks. Id. at 30. And fourth, and perhaps most troubling, there are "no standards in the field whatsoever" for differentiating class and sub-class from individual characteristics. Id. at 31.

With respect to the examiner's failure to take measurements and record those measurements, Lamagna explained: "That's one of the basic violations of the scientific method, if you cannot properly reproduce your measurements so that some other, real scientist can understand what you've done, then you've violated one of those basic elements of the scientific method." Id. at 114.

In addition, Lamagna cast doubt on O'Shea's willingness to declare a match "to the exclusion of all other guns in the world." As Lamagna explained: "No responsible scientist would make such a statement." Id. at 55. "The legitimate way to render an opinion is a statistical opinion. The probability of a match, not 'I've made an absolute match to the exclusion of all other firearms just because I find a few matching striations,' for example." Id. at 106.[25]

III. LEGAL FRAMEWORK

The United States Supreme Court's holdings in Daubert v. Merrell Dow Pharmaceuticals, Inc., 509 U.S. 579 (1993) and Kumho Tire Co. v. Carmichael, 526 U.S. 137 (1999) derive from the Court's concern about the impact of expert testimony on the jury. The rules of evidence give expert witnesses greater latitude than that afforded to other witnesses. See Fed. R. Evid. 703. In addition, a certain patina attaches to the testimony, running the risk that the jury, labeling it "scientific," will give it more credence than it deserves. These concerns are especially present in the case at bar, where the ballistics testimony purports to conclude that the Hi Point weapon found at 6 Esmond Street was "the one" that fired the

[25] Lamagna also suggested available technology that would improve toolmark analysis. The optical, comparison microscope that O'Shea used is "very limited" because it only shows two dimensions. (Daubert Hr'g Tr. 33, Nov. 21, 2005.) To improve the examination, Lamagna recommended the use of a white light inferometer or a stereo or electron-scanning microscope, to better appreciate the contours of the striae -- their three-dimensional topography, as well as their length and width. Id. at 56-57, 63-64. A stereo microscope also provides better control of the light, id. at 59-60, which would avoid the problematic glare featured so prominently in O'Shea's photographs.

shells found a year earlier, “to the exclusion of all other weapons in the world.”

The Court is charged with reviewing expert testimony not just to evaluate its relevancy, but also to determine if it meets the additional standards of reliability under Daubert, Kumho, and Federal Rule of Evidence 702. The Court must find that “the expert’s conclusion has been arrived at in a scientifically sound and methodologically reliable fashion.” United States v. Mooney, 315 F.3d 54, 63 (1st Cir. 2002). The government -- the proponent of the evidence -- bears the burden of proof by a fair preponderance of the evidence. Daubert, 509 U.S. at 593 n.10 (holding that “[t]hese matters should be established by a preponderance of proof”).

In Daubert, the Court articulated four non-exclusive factors that bear on the question of scientific soundness and reliability: 1) whether the expert’s technique or theory can be or has been tested; 2) whether the expert’s technique or theory has been subject to peer review and publication; 3) the known or potential error rate; 4) the existence and maintenance of standards and controls; and 5) whether the technique has gained general acceptance in the relevant scientific community. Daubert, 509 U.S. at 593-94.

The Supreme Court extended the Daubert standard to cover “all expert testimony,” including fields that are not traditional scientific fields, as in the case at bar. Kumho Tire, 526 U.S. at 147. Federal Rule of Evidence 702 was

amended to reflect the Court's decision in Daubert. The Rule now reads:

> If scientific, technical, or other specialized knowledge will assist the trier of fact to understand the evidence or to determine a fact in issue, a witness qualified as an expert by knowledge, skill, experience, training, or education, may testify thereto in the form of an opinion or otherwise, if (1) the testimony is based upon sufficient facts or data, (2) the testimony is the product of reliable principles and methods, and (3) the witness has applied the principles and methods reliably to the facts of the case.

Fed. R. Evid. 702.

Several general principles of the Daubert-Kumho line should be noted: First, Daubert both raised and lowered the standard for the admission of expert testimony. It lowered the standard because it made it easier to admit evidence based on new scientific theories not yet generally accepted in the field, albeit after screening by the court. Hines, 55 F. Supp. 2d at 66. At the same time, however, Daubert plainly raised the standard for existing, established fields, "inviting a reexamination even of 'generally accepted' venerable, technical fields." Id. at 67. Refusing to do so would be equivalent to "grandfathering old irrationality." Id. at 68 n.13. Plainly, I am obliged to critically evaluate toolmark and ballistics[26] evidence, even though it has been accepted for years pre-Kumho.

[26] Ballistics is the study of flying projectiles, including bullets. Toolmark analysis, the technique used in this case, involves the study of marks made by tools, such as the marks a gun imprints on bullets or shell casings. (Daubert Hr'g Tr. 6-7, Nov. 21, 2005.)

Sergeant O'Shea claims to be able to compare shell casings to determine if they came from the same gun, and then to compare those casings with a gun found a year later to determine if that gun had fired those shells. While this is not traditional science, it does not mean it is without scientific pretension: It is empirical, based on observations of physical objects. It makes assumptions about the physical world: 1) that each gun -- like individual DNA -- is unique, because it is made by a metal tool that changes over time; 2) the use of the gun by the consumer causes it to wear in a unique way; 3) the gun's unique signature will be transferred to the projectiles that emerge from it, imprinted on them through the firing pin; 4) an expert can identify that unique signature by visual comparison. There is no reason why these premises and observations cannot be tested under the Daubert-Kumho standards -- using sound research methods yielding meaningful data on error rates. The problem is that they have never been tested in the field in general, or in this case in particular.

Second, the Daubert-Kumho analysis is tempered by the setting, that this is a court, not a scientific conference. In Daubert, the Court noted the difference between information gleaned in a scientific setting and information presented in a courtroom. I have to make the threshold determination under Rule 702 about whether the "expert's conclusion has been arrived at in a scientifically sound and methodologically

reliable fashion." Ruiz-Trioche v. Pepsi Cola of Puerto Rico Bottling Co., 161 F.3d 77, 85 (1st Cir. 1998)(quoting Daubert, 509 U.S. at 590). And in making that decision, I am also to evaluate the evidence in light of the jury's unique role; I assess "not just how valid the data is, but how well the jury can understand it after direct and cross examination, and legal instructions." Hines, 55 F. Supp. 2d at 65.

In other words, the expert's methods must be evaluated, not only for my gatekeeping role, but also to understand the impact of the evidence on the jury's job as the factfinder. The issue is whether "(a) the opinions and conclusions of the expert are accompanied by information that enables the factfinder to evaluate the likely accuracy of the expert's opinion, and (b) the information is presented in such a way that factfinders will not be fooled into excessively overvaluing the testimony." Michael Saks, The Legal and Scientific Evaluation of Forensic Science (Especially Fingerprint Expert Testimony), 33 Seton Hall L. Rev. 1167, 1167 (2003) [hereinafter "Saks, Forensic Science"]. Expert evidence should not be excluded merely because witnesses practicing in that field make errors with some frequency, id. at 1168, but also because the factfinder has no information about the likelihood of error in the opinions, and thus cannot adjust the weight to be given to the evidence.

IV. THE CASE AT BAR

A. Reliability and Testing

The question of whether the expert's technique or theory is scientifically reliable is a specific one: The issue is not whether the field in general uses a reliable methodology, but the reliability of the expert's methodology *in the case at bar*, i.e. whether it is valid for the purposes for which it is being offered, or what the Court has described as a question of "fit." Daubert, 509 U.S. at 591 (quoting United States v. Downing, 753 F.2d 1224, 1242 (3d Cir. 1985)). It may well be that each firearm produces a unique signature transferred onto a shell case and that it is possible to identify that signature using scientifically valid methods. The question is whether the approach used by the expert in this case allows for that identification "to the exclusion of every other firearm in the world."

In General Electric Co. v. Joiner, for example, the court noted:

> Trained experts commonly extrapolate from existing data. But nothing in either Daubert or the Federal Rules of Evidence requires a district court to admit opinion evidence which is connected to existing data only by the *ipse dixit* of the expert. A court may conclude that there is simply too great an analytical gap between the data and the opinion proffered.[27]

[27] In Kumho Tire, the Court was not addressing the reasonableness in general of a tire expert's use of a visual and tactile inspection. Instead, it addressed the reasonableness of such an approach, along with the expert's method of analyzing the specific data he obtained, to draw a conclusion in the case at bar. The question was whether *this* expert could reliably determine

522 U.S. 136, 146 (1997).

Defendants suggest that the contested conclusion in this case -- a match to the exclusion of "every firearm in the world" -- is too great a leap from O'Shea's data. First, they point to the fact that O'Shea was given a single firearm, under circumstances that strongly suggested it was the incriminating weapon, equivalent to an evidentiary show-up, not a line-up.[28] As the Supreme Court noted in Manson v. Brathwaite, a show-up raises reliability concerns because it is a "suggestive procedure." 432 U.S. 98, 107 (1976). In contrast:

> In an evidence lineup, the examiner would be presented with multiple specimens, some of which were 'foils.' The examiner would, of course, be blind to which items of evidence in the evidence lineup are foils and which are the true questioned evidence. For example, a firearms examiner might be presented with a crime scene bullet and five questioned bullets labeled merely 'A' through 'E.' Four of those bullets will have been prepared for examination by having been fired through the same make and model as the crime scene bullet and the suspect's bullet had been. The task of the examiner would then be to choose which, if any, of the questioned bullets were fired through the same weapon as the crime scene bullet had been.

Risinger, Observer Effects, 90 Cal. L. Rev. at 48.

---

the cause of *this* tire's separation. 526 U.S. at 158.

[28] The concern is with "observer bias." It is the reason that there are double-blind studies and placebos, and professors grade exams without knowing the identify of the students. See generally D. Michael Risinger et al., The Daubert/Kumho Implications of Observer Effects in Forensic Science: Hidden Problems of Expectation and Suggestion, 90 Cal. L. Rev. 1 (2002) [hereinafter "Risinger, Observer Effects"].

Second, they note that the absence of notes and photographs in the initial examination makes it difficult, if not impossible, for another expert to reproduce what O'Shea did, something with which Judge Saris was especially concerned in United States v. Monteiro. United States v. Monteiro, No. 03-10329, slip op. at 9-11 (D. Mass. Nov. 28, 2005). Reproducibility is an essential component of scientific reliability.

The fact that O'Shea reconstructed the original examination five years later raises even more concerns.

> [C]onsider 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.

Id. at 15-16.

Here, O'Shea reexamined the evidence and photographed some of it in preparation for a hearing in a particular criminal case, knowing that the defendants had been indicted for these crimes, knowing that the trial was about to begin, that his examinations were being questioned, and that he would shortly be testifying about his original examination. Nothing would prevent the witness from identifying one set of marks as

similar in the 2001 examination and yet another on reexamination, without notes from the first. Neither defense counsel nor the jury would be in a position to evaluate just how subjective and standardless the methodology was.

Finally, O'Shea had no coherent database with which to compare the shell casings he was reviewing -- neither a national database identifying the class and sub-class characteristics of particular firearms or his own database from his experience with other weapons. At the time of his examination, he did not call the manufacturer of this Hi Point weapon to ask what the 1998 production process was. The most he did was look at shell casings from other weapons (which he knew were not involved in this case) to determine class and sub-class characteristics. Without knowing when these guns were manufactured, the comparison could only be at the most general level.[29] Nor did O'Shea rely on national standards or even Boston Police protocols. There were none in 2001.[30]

Plainly, these issues -- the reliability of the methods used and whether they can be tested -- point against the admission of the testimony.

---

[29] In 2002, the Boston Globe reported on efforts to create a ballistic database to record the unique characteristics of shots fired by each newly manufactured weapon. While Boston police officials, lawmakers, and handgun control advocates called for such a database, the article reported that it was opposed by "gun advocates." John Ellement, Call Rises to Create Ballistic Database Gun 'Fingerprint' Called Police Tool, Boston Globe, Oct. 24, 2002, at B12.

[30] Even in fingerprint analysis, defendant noted that there was a "one-dissimilarity doctrine." Under this approach, if two fingerprints contain one genuine dissimilarity, the prints may not be attributed to the same finger or individual.

B. Error Rates and Proficiency Testing

Even if O'Shea's approach did not account for observer bias, the Court and ultimately the jury could still evaluate the testimony by considering the error rates in the field and the error rates for this examiner. In other words, even if his approach may be flawed, if examiners in the field manage to overcome those flaws, or if this examiner had a low error rate, the evidence may still be reliable, and the jury can evaluate it. See Kumho Tire, 526 U.S. at 152. Without information about error rates, the initial factfinder, this Court, and the ultimate one, the jury, have no accurate way of evaluating the testimony.[31]

Here, there was no credible testimony about the error rates of this examiner or in the field as a whole. O'Shea could not say that his work, his approach, his conclusions were certified by any neutral body to assure that he passed minimal standards. The Court had to rely solely on his testimony that he had examined hundreds of casings and guns. In effect, the jury would have to trust in his observational capacities, without knowing how often he was actually correct.[32]

---

[31] As Professor Saks explains, the question is: "When people in your field offer opinions regarding this task, how accurate are they?" If the expert could not give an informative answer to such a question based on sound and adequate data -- that is, if the expert's honest answer would have to be "I don't know" -- then the testimony is not helpful to the jury and is vulnerable to exclusion. Saks, Forensic Science, 33 Seton Hall L. Rev. at 1170.

[32] Nor is there evidence of any peer-reviewed publications in the ballistics/toolmark field as that idea is understood in Daubert and Kumho.

These factors, the absence of testimony about error rates, certification, or proficiency testing, point against the admission of the testimony.

C. Significance of Defense Expert and Availability of Cross-Examination

Even without testimony about error rates, proficiency testing, and certification, I could conclude, as I did in Hines, that if the jury is able to see and understand what the expert saw, then the testimony may be admissible. If the jurors cannot see and understand the testimony, it amounts to nothing more than "trust me" testimony, the kind of *ipse dixit* with which the court was concerned in Joiner.

The government maintains that O'Shea's testimony is readily accessible to the jury, that the points the defense wishes to make about subjective testing, error rates, and other methodological weaknesses can be easily understood by jurors. Moreover, since there was no destructive testing in the case at bar, the defense has its own expert in a position to review the evidence. The issues are not so complex, not so technical, that the jury will not understand.

This position does not completely address the threshold problem: the reliability of the testimony. Obviously, if the testimony is wholly unreliable, so as not to meet the Daubert-Kumho threshold, it would not matter how well the jury would understand it. Nevertheless, on balance, this factor tilts *in favor* of admission, so long as the Hines limits are in place.

D. Significance of Precedent

In its opposition to defendant's motion, the government notes that the defendant "cannot cite a single case from any court, state or federal, trial or appellate, in which ballistics evidence of the same type at issue in this case was deemed unreliable." Def. Opp. at 7. There is apparently widespread acceptance in the courts of ballistics testing and toolmark analysis. True enough. Although the scholarly literature is extraordinarily critical,[33] court after court has continued to allow the admission of this testimony.

Several court opinions rely upon the long-standing recognition of ballistics evidence in courts. The Fifth Circuit, for example, has found that "the matching of spent shell casings to the weapon that fired them has been a recognized method of ballistics testing in this circuit for decades." United States v. Hicks, 389 F.3d 514, 526 (5th Cir. 2004). That court further found that "[b]ased on the widespread acceptance of firearms comparison testing, the existence of standards governing such testing, and [an expert's] testimony about the negligible rate of error for

[33] See I.A. Pretty & D. Sweet, The Scientific Basis for Human Bitemark Analyses-A Critical Review, 41 Sci. & Just. 85, 86 (2001) ("Despite the continued acceptance of bitemark evidence in European, Oceanic and North American Courts, the fundamental scientific basis for bitemark analysis has never been established."); Craig M. Cooley, Reforming the Forensic Science Community to Avoid the Ultimate Injustice, 15 Stan. L. & Pol'y Rev. 381, 382 (2004) ("It seems that the only standard the courts are requiring of forensic science is that it be incriminating to the defendant." (footnote omitted)); Adina Schwartz, A Systemic Challenge to the Reliability and Admissibility of Firearms and Toolmark Identification, 6 Colum. Sci. & Tech. L. Rev. 2 (2004-2005) (contending that despite widespread faith in "ballistics fingerprinting," firearms and toolmark identifications should be inadmissible across-the-board because of systemic scientific problems).

comparison tests, the district court had sufficient evidence to find that [the expert's] methodology was reliable." Id. In a death penalty case in the District of Maryland, the court rejected a Daubert challenge to ballistics evidence. In that case, the defense did not offer contrary testimony to refute the government's ballistics expert. The court wrote, "Ballistics evidence has been accepted in criminal cases for many years. . . . In the years since Daubert, numerous cases have confirmed the reliability of ballistics identification." United States v. Foster, 300 F. Supp. 2d 375, 377 n.1 (D. Md. 2004).

Even the Supreme Court has weighed in on ballistics evidence. In United States v. Scheffer, the Court contrasted polygraph evidence with more acceptable forms of expert testimony, including ballistics, in which the expert "testif[ies] about factual matters outside the jurors' knowledge." 523 U.S. 303, 313 (1998). While this was merely a casual reference to ballistics, likely without any argument on the issue, many lower courts have cited this opinion as validating the use of ballistics experts. See, e.g., Foster, 300 F. Supp. 2d at 377 n.1; U.S. v. Williams, 2004 U.S. Dist. LEXIS 25644 (S.D.N.Y. 2004).

State courts have similarly rejected Daubert-type challenges to ballistics testimony. See, e.g., State v. Brewer, 2005 Conn. Super. LEXIS 806, 3 (2005) (finding that "[t]he testimony of the state's expert with regard to firearms

and ballistics is so well established that it does not require analysis" under the state's Daubert-type rule).

This reliance on long-standing use of ballistics evidence in the courts is troubling. It runs the risk of "grandfathering in irrationality," without reexamining it in the light of Kumho and Daubert. It arguably ignores the mandate of Daubert, especially where the courts are relying on pre-Daubert acceptance of a given scientific technique.

The First Circuit has not addressed the issue, although it has dealt with related forensic fields. In Hines, this Court chose to allow a handwriting expert to testify to similarities in handwriting but not to testify to the ultimate conclusion that two samples matched. Hines, 55 F. Supp. 2d 62. Another district judge, however, declined to follow this same rationale and allowed a handwriting expert to testify to his opinion that two samples matched.[34] Affirming this decision to admit the expert testimony, the First Circuit found no abuse of discretion in the judge's explanation that "the reliability of the handwriting comparison testimony and the expert's ultimate opinion on authorship were inevitably linked because they were based on the same methodology." United States v. Mooney, 315 F.3d 54, 63 (1st Cir. 2002). Mooney does not necessarily mean that such evidence must be admitted; rather, it suggests that

[34] The record reflects that the court heard argument on the motion before trial; it is not at all clear whether or not the court heard testimony from experts pro and con, as this Court had done.

the trial court has discretion to either include or exclude expert testimony in this context.

This precedent plainly points *in favor* of admissibility.

V. CONCLUSION

Putting together this precedent with the evidence I have heard, suggests admission but with limitations, limitations identical to those I adopted in Hines. O'Shea is a seasoned observer of firearms and toolmarks; he may be able to identify marks that a lay observer would not. But while I will allow O'Shea to testify as to his observations, I will not allow him to conclude that the match he found by dint of the specific methodology he used permits "the exclusion of all other guns" as the source of the shell casings. Defense will be permitted full and fair cross-examination.

I therefore **GRANT IN PART** and **DENY IN PART** Defendants' Motion to Exclude Ballistics Evidence [document #434].

SO ORDERED.

Date: December 20, 2005 **/s/NANCY GERTNER, U.S.D.J.**