Review
AB1717 report
Technical Evaluation: Feasibility of a Ballistics Imaging Database for All New Handgun Sales

Review
AB1717 report
Technical Evaluation: Feasibility of a Ballistics Imaging Database for All New Handgun Sales

Dr. Jan De Kinder
Ballistics Section, Head
National Institute for Forensic Science (NICC/INCC)
Department of Justice
Vilvoordsesteenweg 98-100
B-1120 Brussels
Belgium

jan.dekinder@just.fgov.be

I, hereby, declare that during this study I was not influenced by either party who collaborated to this study and the rebuttals.

Executive Summary

This is an independent review of the project report entitled "Technical Evaluation: Feasibility of a Ballistics Imaging Database for All New Handgun Sales" (referred to hereinafter as the "AB1717 Evaluation" or the "Evaluation"). The experimental work conducted in the Evaluation was found to be in agreement with general scientific principles and with the current knowledge in the field of forensic firearms identification. Rebuttals to the project report from the Bureau of Alcohol, Tobacco, and Firearms (BATF), and by the system manufacturer, Forensic Technology Inc. (FTI), were examined and dismissed for the reasons detailed below.

A ballistic imaging database system is based on the following premise: When a gun is fired, it may leave distinguishing marks on the fired bullet and cartridge case. A searchable database with images of such marks from all guns sold could be a valuable investigative tool. Such a database would permit linking of evidence bullets or cartridge cases back to the gun that fired them. Evidence found at a shooting incident would be scanned and compared to all entries in the database. Ideally, the system would provide investigators with the serial number of the firearm. The serial number can lead the investigators to the registered owner. The database could be called a "ballistic fingerprinting" database. If created, California's ballistic fingerprinting database would quickly grow to be very large. More than 100,000 fired cartridges from new pistols would be added to it annually.

No technology designed specifically for ballistic fingerprinting exists for the moment. The purpose of the AB1717 Evaluation was to determine whether or not the Integrated Ballistic Identification System (IBIS™) could be adapted successfully for this new role. While similar in concept, IBIS™ is used successfully with numerous regional but much smaller "open case file" databases. IBIS™ performs automated comparisons between bullets and cartridge cases from different crime scenes. IBIS™ is the cornerstone of the National Integrated Ballistics Information Network (NIBIN), deployed by BATF.

The AB1717 Evaluation was designed to test the performance of the IBIS™ system for the anticipated large database of new firearms. The experiment used 792 Smith & Wesson model 4006 semi-automatic pistols for this purpose. Each pistol was test fired using at least two cartridges of Federal brand ammunition and other ammunition. One of the test fired Federal cartridge cases for each of the pistols was registered into the database.

The duplicate Federal cartridge cases from fifty of these pistols were selected at random and compared with the database. The system ranks how well each entered mark matches the evidence. The higher the ranking the more similar the stored image is to the evidence's mark. For the system to be successful, the correct gun should be listed in the top few ranks. The results show that 38 % of the fifty pistols were not listed in the top 15 ranks. The same experiments was repeated with ammunition of a different brands. In this case 62.5 % of the pistols were missed and not listed in the top 15 ranks. These results will be discussed in light of the investment in terms of equipment and personnel needed to set up a ballistic fingerprinting database. In fact, the trends in the obtained results show that the situation worsens as the number of firearms in the database is increased.

Two rebuttals were received. One was from the BATF and the other from Forensic Technology Inc., the manufacturer of the IBIS™ system. The arguments in both rebuttals are discussed in this review. The main argument of both rebuttals is as follows:

(1) From the BATF: Federal ammunition has primers that are too hard. The BATF argues that this may have skewed the results. They prescribe Remington-Peters ammunition for test firing into the IBIS™ systems. Their choice is not based on peer-reviewed published research. In fact, hardness measurements show that the primers of

Federal cartridges have a lower hardness value than Remington-Peters, contrary to the BATF report. The BATF fails to see that other parameters of the cartridge also play a very important role in the marking of the primer. Taken into account all parameters, Federal has fair-to-good performance and was an acceptable choice for the study.

(2) From FTI: Out of the 50 duplicate cartridge cases used in the Evaluation, eight could not be matched through manual examination by one of their firearms examiners. FTI proposed to remove them from the statistics to achieve better results. This is unacceptable. As the AB1717 Evaluation discusses the applicability of an automated comparison system to the problem of mass-produced firearms, all data points have to be taken into consideration. The goal of a ballistic fingerprinting system is not restricted to those cartridge cases that can be identified by a trained firearm examiner.

Based on the results of this test, a number of methods can be developed that may increase the performance of the IBIS™ automated correlation system when applied to the problem of ballistic fingerprinting. Different protocols for the collection of the reference material should be developed and tested to optimize the correlation performance of the ballistic fingerprinting. This can be done by varying (at least) two parameters:

- Optimization of the choice of ammunition used to test-fire the guns. By correlating different types of ammunition fired through a single firearm, the best suitable ammunition for test firing can be determined.
- Multiple images can be entered per firearm. This allows for a certain variation in the marks left on different brands of cartridge cases. In a previous study, this protocol increased the correlation performance strongly.

Apart from this, Forensic Technology can try to improve the performance of the technology. This can be done in several ways suggested in the paper. FTI's suggestion to start a pilot project should also be considered. Similar results can be obtained by monitoring the results from the ballistic fingerprinting programs in Maryland and New York.

It is important to mention that when starting a ballistic fingerprinting database, the technology and the protocols have to be well established and oriented towards future compatibility. A evaluation of different technologies has to be performed, prior to choosing for an existing solution. If this is not done so, the chances are that the now established database will be rendered obsolete in a couple of years.

Introduction

Constraints

This independent review of the results from the project entitled 'Technical Evaluation: Feasibility of a Ballistics Imaging Database for All New Handgun Sales' (referred to hereinafter as the 'AB1717 Evaluation') envisages only the scientific and technical aspects relating to the forensic problem. Other aspects such as legality, allegations of partiality, the amount of necessary skilled personnel, and the practical arrangements of collecting and storing the necessary material are not addressed in this document. I refer the interested reader to the paper entitled "Ballistic Fingerprinting Databases[1]" on this subject.

Consulted information

The following documents were examined in this study:

Document	Ref.
AB1717 Evaluation by the Bureau of Forensic Science (F. Tulleners et al.)	A
Forensic Technology Rebuttal of June 25, 2001	B
Forensic Technology Rebuttal of March 18, 2002	C
- Exhibit A: The Missing Link: Ballistics Technology that Helps Solve Crimes	D
- Exhibit B: Glock Cartridge Cases Acquisition and Correlation Test Protocol and Results + Presentation of the Glock Results to the CAL DOJ Stakeholders Meeting	E
- Exhibit C: The Methods and Technology for 'Ballistic Fingerprinting' and their Practical Applications	F
BATF Rebuttal of May 13, 2002	G
BATF cited research articles	
- R. Tontarski and R. Thompson, J. Forensic Sci. **43(3)**, 641-647 (1998)	H
- R. Thompson, M. Desrosiers, and S. Hester, AFTE Journal **28(3)**, 194-203 (1996)	I
- R. Thompson, abstract AAFS meeting, San Francisco (1998)	J
- J. Miller and M. Mclean, AFTE Journal **30(1)**, 15-61 (1998)	K
- P. Lardizabal, AFTE Journal **27(1)**, 49-51 (1995)	L
- E. Thompson, AFTE Journal **28(2)**, 95-96 (1996)	M
- J. Miller, AFTE Journal **30(4)**, 631-638 (1998)	N
- J. Miller, AFTE Journal **32(3)**, 259-270 (2000)	O
- H. Silverwater and A. Koffman, AFTE Journal **32(1)**, 32-39 (2000)	P
BATF cold hit definitions April 2002	Q
BATF Protocols 1995 and 1999	R
Tulleners Response memo to Forensic Technology Rebuttal of June 25, 2001	S
Lansing Lee's Response memo to BATF Rebuttal of May 13, 2002	T
Excel spread sheet showing the top ten-cartridge case breech face hits for different ammunition - Federal, Remington, and Winchester etc.	U
Tulleners Response Memo to the BATF Rebuttal of May 13, 2002	V
Protocols for the proposed 9 mm Sig Sauer Experiments	W
Crime Gun Trace Reports (2000)[2] - BATF	X
Primer hardness tests performed by Dr. Eric Randich (Lawrence Livermore Natl, Lab)	Y

[1] Ballistic fingerprinting databases - Jan De Kinder, scheduled for publication in Science and Justice, Vol. 42 (2002).
[2] Document downloaded from the BATF website, http://www.atf.treas.gov/firearms/ycgii/2000/index.htm

Reference to page 3 of a cited document A is indicated as [A, p.3].

Organization of this paper

Numerous remarks were made in both rebuttals from Forensic Technology Inc. (FTI) as well as from the Bureau of Alcohol, Tobacco and Firearms (BATF) and the response memo's to these documents. Some of the remarks are very important to the scientific discussion. Others are relevant, but appear on a second or third layer of interest.

In order to guide the reader through this complicated and technical matter and not become entangled in minutia, this review is organized as follows. In the second chapter, a critical and independent review of the design of the project is performed taking into account the major remarks from both rebuttals. The results from the eight performance tests are discussed. The following chapter deals with arguments of secondary order made in at least one of the rebuttals. This documents ends with a number of suggestions for future research.

1. Design of the project

The AB1717 Evaluation aims at applying the currently available automated comparison systems to the concept of mass sampling of manufactured firearms (setting up a ballistic fingerprinting database) for the State of California. The discriminative power of the searching algorithm of such a database was tested for a large number of firearms of a single model. The marks left by them on a cartridge case were input into a database. A number of questions originating from the field of forensic firearms investigation were investigated:

- Is the system capable of identifying cartridge cases back to the firearm ?
- What if ammunition of different brands is used ?
- What is the time necessary to obtain results out of such a system ?
- Is there any effect from the aging or use of the firearm ?
- Can the ballistic signature be easily modified to mislead the system ?

1.1 Firearms

In a first step, attention is focused on semi-automatic pistols, as they are most frequently associated with crime. In this study, 792 new semi-automatic pistols, manufactured by Smith & Wesson, model 4006 of caliber .40 Smith & Wesson were made available by the California Highway Patrol. This is a reasonable choice, given the availability of this firearm and the convenient test firing facility adjacent to the firearms storage location. No objections were made in the rebuttals.

It seems to me that a potentially important information is missing: the report does not mention explicitly that the class characteristics of all the 792 firearms were the same. No information on their serial numbers is available. This could provide interesting forensic information if the firearms were manufactured consecutively.[3]

The caliber which was the most frequently used in crime in 2000 for the state of California is the 9 mm Luger [A, p.5-3], accounting for 40%. It would be logical to use this caliber for testing. The occurrence of the caliber .40 Smith & Wesson was also substantial at 11 % [A, p.5-3]. In 2001, 234 Smith & Wesson, model 4006 firearms were sold in California. The total number sold designated as .40 Smith & Wesson was 20,135.[4] They all have the same class characteristics that might be imparted unto cartridge cases.[5]

An important parameter in the discussion of this type of databases is the time between the selling of a firearm and actually committing a crime with it. This so-called time to crime (TTC) is discussed for California in Section 3.3. This parameter allows one to minimize the retention time of the data in the ballistic fingerprinting database without losing a substantial amount of performance. The full size of the database as well as the amount of firearms of one type can be estimated.

In the AB1717 Evaluation [A, p.1-3], a retention time of 5 years is proposed, which seems to be illogical, compared to the average TTC for California, which is higher than 6.1 years. For the ballistic fingerprinting database to be sufficiently successful, at least a retention period of

[3] A list of all the serial numbers was passed to me by Fred Tulleners by email on 26.11.02 as reply onto draft version of this review. Its analysis is beyond the scope of this document.
[4] Figures passed to me by John Rush (California Department of Justice) by email on 31.10.02 after having requested this information.
[5] Unchecked information from John Rush (California Department of Justice) by email on 30.11.02 as reply onto draft version of this review. The NICC/INCC has insufficient pistols of this brand/model to verify his statement.

15 years should be applied. This strongly enlarges the estimated size of the database. Moreover, the discussion on when a firearm is “home free” remains [A, p.5-5]. We can expect to have at least 3,300 Smith & Wesson, model 4006 firearms entered in the database.

The number of 792 firearms seems to me a good starting point for this research plan.

1.2 Ammunition

Ammunition manufactured by Federal was chosen to perform the test-firings. This selection was predicated on the fact that only one vendor had sufficient ammunition from *one* lot during the initial purchase time frame.

An important argument in the BATF rebuttal concerns the ammunition used. The BATF rebuttal states that [G, p.15-16] "Unfortunately, Federal brand ammunition was used in the research described in the Evaluation. (…) Federal ammunition is not prescribed by the BATF protocol ammunition in any of the calibers of interest, due to the primer surface generally being too hard in comparison to the ammunition being used in handguns.(…) If protocol ammunition had been used, it is likely that the match/ranking results would have been much higher."

A lot of weight was given to this argument in the BATF rebuttal, given the peremptoriness of the words used: "There is a glaring methodology flaw in the study design that colors the whole study, the data from that method, and necessarily the purported results of the data. (…) The results of correlations, the determination of drop-out of candidates in a growing database, the ability for this reference set of casings to find other brand matching casings - all of these results are skewed due to the selection of Federal Brand ammunition." [G, p.15]

Fig 1 Comparison between two cartridge cases of the same brand, from different batches (carrying a different headstamp) fired through the *same* firearm.

Primer hardness

The literature cited at the end of their rebuttal does not include a controlled variation of ammunition brands to support the BATF-protocol[6]. In general, I have not found any information on the hardness of primers in peer-reviewed literature. The manufacture of primers is, in general, a process with very high quality control. The primer metal has to meet a specification as to content and hardness. The variation of these characteristics is beyond the control of the forensic firearm examiner. However, abrupt changes may occur. My laboratory recently had the experience of not being able to link two known matches using two different lots of ammunition (Sellier & Bellot) on basis of the breech face impression. (see Fig. 1) These rounds carried a slightly different head stamp, indicating that they came from different lots of ammunition, while all other characteristics remained the same.

Hardness measurements were performed by Dr. Eric Randich at the Lawrence Livermore National Laboratory on six different brands of .40 Smith & Wesson ammunition used in the Performance Test 3.[Y][7] Table 1 shows the results for cartridges of caliber .40 Smith & Wesson.

Table 1 Hardness measurements on primers (results from the LLNL).

Manufacturer	Lot No.	Primer Type	Hardness (Vickers)
Remington Peters	H29NC2517	Nickel	157 +/- 12 Hv
Federal	420322X269	Brass	108 +/- 5 Hv
Winchester	RC41	Brass	114 +/- 6 Hv
Armscor	03093000	Nickel	159 +/- 18 Hv
Corbon	40SFA050	Nickel	186 +/- 13 Hv
Eldorado (PMC)	RC1078	Nickel	166 +/- 14 Hv

Three indentation tests were performed on each primer for each manufacturer. Quite remarkable, Federal seems to have the lowest hardness for the primer material.

Protocols

When identifying a firearm to recovered bullets and cartridge cases for a crime scene, ammunition is selected for test-firing in the following order (from most to least suitable):

- ammunition from the same lot as the recovered bullets and cartridge cases,
- ammunition from the same type as the recovered bullets and cartridge cases,
- ammunition from the same manufacturer as the recovered bullets and cartridge cases, having the same primer composition and/or bullet jacket composition,
- ammunition having the same primer composition and/or bullet jacket composition.

This means that the results between Federal and Federal cartridge cases (Performance Test 1) are in full agreement with commonly accepted practice.

The common forensic experience applies to situations in which the bullets and cartridge cases originating from the crime scene were seized prior to test firing the weapon. In setting up a ballistic fingerprinting database, the situation is reversed and the above mentioned

[6] Even more remarkably: in R. Thompson, M. Desrosiers, and S. Hester, AFTE Journal **28(3)**, 194-203 (1996), which is referenced as [I], Federal ammunition was used to perform the tests.
[7] email from Fred Tulleners, dated November 12th, 2002.

standard procedure cannot be applied. One does not know in advance the brand of the ammunition to be found on the crime scene. Moreover, it can be a brand which will only become available on the market in the future or it may bear characteristics which are not consistent with the ammunition used to test fire the weapon.

Correlation tests

Whereas the hardness is an important parameter for its acceptance of breech face markings, other factors such as primer seating, building up of the gas pressure, etc… may be even more important. A test aimed to determine the mutual influence of the 18 most common primers was done by the Forensic Institute in The Netherlands[8]. This tests accounts for all the contributing parameters. Unfortunately, Remington ammunition was not included, as it is very uncommon in Europe. Cartridges from different brands were fired in the same gun and correlated by the IBIS-system. The results for a number of brands are shown in Table 2. The abbreviated brand name is given as well as the composition of the primer, brass (B) or nickel (N). The results show that Federal has a medium performance. The best brand was Winchester with a nickel primer. Bad marking occurs for Hirtenberger, Fabrique Nationale and Norma. The results of this test are only indicative, as the IBIS-system has substantially evolved over the last 6 years.

Table 2 Matrix of the highest correlation results of primer tests. For the gray cells, the highest score is obtained on the breech face marks. For the white cells, the highest score is on the firing pin impression.

	ACE-N	DAG-N	GECO-N	GFL-N	NORMA-N	PMC-N	PPY-N	SILVA-N	WIN-N	FC-B	FNB-B	GECO-B	GFL-B	HP-B	MRP-B	NORINCO-B	S&B-B	SPEER-B
ACE-N		3	5	7	11	3	4	7	1	6	5	1	15		4	1	1	5
DAG-N	3		23	5	2	2		12	2	2		3	1			18	8	4
GECO-N	7	14		11		5	22	10	6	1	4		22		2	9	2	2
GFL-N	2	1	5		34	1	1	1	1		1	1		1	2	27	3	3
NORMA-N		3		14		10						7	4					20
PMC-N	1	2	3	2	7		5	4	3	1	2	3	6		5	2	6	1
PPY-N	4		31	4		6		2	9		2	5	3	2	8	3		
SILVA-N	5		3	1			2		3		9		9	5	3			
WIN-N	3	1	2	3	2	2	8	2		5	1	2	45	3	7	6	4	4
FC-B	28	8	6	23		8	14	18	20		16	33	27		9	11	9	8
FNB-B			15			14	6	4						4	6		35	5
GECO-B	2	5	10	3	4	5	3	6	5	8	6		2			2	1	3
GFL-B	10	6	28	7	8	16		34	4	12		19			16	4		10
HP-B							24											
MRP-B			20			7	7	3	4	6	3			2			20	1
NORINCO-B	9	13		12	1	11			26	3		10	5				10	13
S&B-B	5	8	4	10		9	25	7	7	4	10	4	13		3	6		6
SPEER-B	22	5	7	8	20	1	10	11	18	7	5	14	16		1	29	7	

[8] IBIS - Evaluation report January 1997 - March 1997, Dutch Forensic Science Laboratory, Weapons and Ammunition Laboratory (1997). The experiments were performed using a IBIS-system version 2.0.

To determine the correct type of ammunition for test firing of the gun prior to its sale, a similar test to the one performed by the Dutch lab should be performed. This test takes into account not only the primer hardness, but also other factors (such as primer seating, the pressure generated by the gun powder, …), which explains why the harder Winchester primers (see Table 1) seem to perform much better.

If one type of ammunition can cover the whole range of currently available brands, it should be selected. If one type of ammunition does not provide sufficiently good correlation output, multiple types of ammunitions have to be selected. Through negotiated supplier contracts, the same characteristics of the primer can be guaranteed. Under these conditions, correct protocols to encode the data can be set up.

A limited analysis can already be performed based on the obtained results listing the top ten breech face and firing pin hits using different makes of ammunition available in the Performance Test 3. [U, p.1] Federal cartridges are used as the reference marks. The results are shown in Table 3.

Table 3 Correlation results for different brands compared to Federal cartridges.

Ammunition Brand	% of top 10 breech face hits	% of top 10 firing pin hits
Federal	38	44
Winchester	38.9	28
Eldorado	36	38
Remington	7.7	23
Corbon	0	0
Armscor	0	7

For a medium hardness primer, good correlation results are expected with almost all types of ammunition. Remarkably, this does not apply for Remington. Similar to the study from the Dutch lab, Winchester turns out to correlate quite well in this limited study.

Different markings of ammunition

The AB1717 Evaluation says that “different types of ammunition can mark differently”. [A, p.1-2]. The BATF-rebuttal states that this is incorrect and that only differences in the depth of marks can be observed on cartridge cases. "While other factors such as the ammunition used can affect the depth to which a firearm makes its marks, such as on a cartridge casing, the marks themselves do not change; rather, the same marks may be shallower on harder cartridge casings and deeper on casings composed of softer metal." [G, p.16] Both opinions can be easily brought together by noting that as the depth of mark decreases, it will become invisible even for microscopic observations.

Conclusion

No valid scientific arguments have been provided by the BATF for its selection of Remington-Peters as the standard ammunition for the protocol. Hardness measurements show that the primers of Federal cartridges have a lower hardness than Remington Peters, contrary to the BATF report. The BATF fails to see that other parameters of the cartridge also play an important role in the marking of the primer. Correlation experiments show that Federal has a

fair-to-good performance. The opinion that "a glaring methodology flaw (…) all of these results are skewed …"[G, p.15] is certainly not correct.

1.3 Automated comparison system

Forensic Technology Inc. is the sole provider of automatic comparison systems for both bullets and cartridge cases. SBC (Russia) sells a system for the automatic comparisons of bullets with a similar discriminative power but providing images of a much higher quality[9]. The product of FTI, Integrated Ballistic Identification System (IBIS), has substantially evolved during the last decade and is the cornerstone of the National Integrated Ballistics Information Network (NIBIN) program. Reference can be made to the FTI rebuttal [C, p.6-7], the BATF rebuttal [G, p.8-13] and to the document entitled "The Missing Link: Ballistics Technology that helps Solve Crime" by the BATF (ATF P 3315.1 (10/01)), referenced as [D].

The product has been developed for aiding laboratories to deal with very large open case files. IBIS allows forensic examiners to compare a newly arrived bullet or cartridge case to a large database of still unsolved shooting incidents in a limited time period. The same applies to test-fired material from seized weapons. The NIBIN program had at the time of the writing of the BATF rebuttal "generated a total of 4429 'hits'. These hits are significant investigative leads for law enforcement authorities to use in fighting violent crime." [G, p.13]

It is important to note that IBIS has not been designed for operating with large databases such as the ballistic fingerprinting database. A minimal requirement for the equipment to be used for ballistic fingerprinting is to be able to deal with open case files. The IBIS system may or may not perform well for mass sampling of manufactured firearms. The content of the AB1717 Evaluation is the main scientific data available so far[10]. In the document "The Methods and Technology for Ballistic Fingerprinting and their Practical Applications", FTI announces a Virtual Serial Number (VSN) system. I do not know if this is operational at this time.

The AB1717 Evaluation turned to the IBIS-system for the automatic comparison to the reference material from the guns in legal possession, as it was the only applicable technology on the market.

1.4 Protocols

I have no remarks on the procedure described to test-fire the pistols [A, p.7-2].

Only the breech face markings and the firing pin impressions of a single test fired cartridge case (randomly selected out of the two Federal cartridges) for every gun were entered into the system.

A single cartridge case

The entry into the system was performed using the same protocol as for entering cartridge cases found at the crime scene. It is my belief that these protocols should be adapted to this new application. A similar flexibility towards protocols can also be seen implicitly in the rebuttal of the BATF [G, p.21], stating that (for the search protocol) "This is not an immutable

[9] V. Thach, "BIS CONDOR", presentation during the 9th meeting of the ENFSI (European Network of Forensic Science Institutes) Firearms Expert Working Group, Bratislava (Slovak Republic) 2002.
[10] Apart from the experiments performed by FTI, using Glock 17 firearms (see Performance Test 1)

characteristic of IBIS, but a protocol developed from experience in using the system and open to change as the system changes. (…) If a situation develops such as the one invoked in the Evaluation, of a very large database with many very similar exhibits, and if in this situation the examination of the top 10 potential[11] high confidence candidates proves to be insufficient, additional images can be examined and the protocol changed accordingly".

A major result from the AB1717 Evaluation (Performance Test 3) is that the marks imprinted on different brands of ammunition do not correlate sufficiently well to go further. A possible solution may be to rely on multiple cartridge cases that are input into the system for a single pistol. If the correlation with a crime scene case is performed, the correlation scores can be combined to provide a more correct result. The article [P] says "The second method requires inputting two cartridge cases, instead of one. Ten case studies showed no missed 'hits', 100 % success". The size of the database is much smaller than for ballistic fingerprinting databases, but the idea is valid and should be explored.

Firing pin imprint and breech face markings

Only the imprint from the firing pin and the breech face of each firearm were encoded into the system. It is not clear to me why the ejector marks were not registered in the system. The reference cartridge cases are pristine, so an analysis of the ejector mark is straightforward.

Three observations can be made :

1. The document entitled 'The Methods and Technology for Ballistic Fingerprinting and their Practical Applications' by FTI (document [F]) specifies that their VSN-system encodes only breech face and firing pin images.
2. FTI personnel encoded the cartridge casings on equipment from FTI.
3. The results by FTI on the Glock firearms were obtained using only the breech face and firing pin impressions.

It is hence strange to notice that FTI remarks that [C, p.6], "In May 1998, IBIS version 3.0 was released. (…) The new version added ejector marks, and is useful for a wide variety of automatic weapons. A significant weakness of the AB1717 Evaluation is that it does not recognize the capability of IBIS in this regard." In their previous rebuttal, they write [B, p.2] "However, IBIS has had strong success with ejector markings also and it is IBIS that is in use in the Ballistic Identification Databank Programs ongoing in New York and Maryland." Please note that the previous sentence does not specify if ejector marks are encoded in the Maryland and New York project.

Taking into consideration the results from this study, it is my belief that a maximum number of marks should be used to enhance the discriminative power of a ballistic fingerprinting database. Further studies should be performed by FTI to evaluate the inclusion of other marks present on fired cartridge cases into their VSN system. [F]

No bullets

Only cartridge cases were entered into the system. It is not mentioned in the report if this was done in order to reduce the time to input the data (the time required to scan bullets is much longer than for cartridge cases) or to be in agreement with the VSN program of FTI.

The report mentions an average time necessary to test fire a pistol and encode a bullet of 84 minutes. In a ballistic fingerprinting system, one can ask the manufacturer to supply a test

[11] The AB1717 Evaluation considers the ranking within the top 15 cartridge cases. It is not clear why the more flexible or conservative figure of 15 was taken.

fired bullet. This reduces the time necessary to encode a pristine bullet to about 20 minutes. However, test fired bullets are more difficult to obtain than cartridge cases. A bullet recovery system, such as a water tank or cotton box is necessary to perform this task. This will necessitate important investments for equipment and personnel.

It is also important to note that the signature from bullets initially develops and becomes only stable after 5-10 rounds of ammunition are fired in the weapon.[12]

1.5 Analysis of the results

Performance test 1 - Results of the 50 Random Cartridge Cases

I fully agree with the analysis of the data as it was performed. "Thus, given the same ammunition for comparison, the system will miss about 38 % of the time when either a breech face or firing pin is used as a ranking score."[A, p.8-4]

A more flexible criterion for finding the corresponding cartridge was used as one looked at the top 15 cartridge cases, whereas the normal procedure only uses the top 10.

FTI objects that "of the 50 duplicate cartridge cases used in the Evaluation, eight could not be matched through manual examination by John O'Neil, a well-known firearms examiner with more than 30 years of experience.(…) Moreover, of the remaining 42 duplicate cartridge casings, approximately half had markings that were somewhat unfavorable. (…) It is quite possible that some of these specimens could not have been matched using microscopes.[C, p.13]

The basic argument from the FTI is that "the performance of an automated examination could not, and should not, be more accurate than a microscope comparison by a firearms examiner". [C, p.14] This is a unscientific argument. It is the same type of expression as saying at the beginning of the 1990's that automated comparison of bullets and cartridge casings is impossible. The correct expression should be that *the current scientific knowledge and state-of-the-art technology does not allow one to be more accurate than a microscope comparison by a firearms examiner.*

FTI proposes to remove them from the statistics to achieve better results: "However, the percentages reported by the AB1717 Evaluation would have been even higher if the unmatchable examples identified by Mr. O'Neil had been eliminated." [C, p.15] In my opinion, this is unacceptable. As the study discusses the applicability of an automated comparison system to the problem of mass manufacturing of firearms, all data points have to be taken into consideration. The goal of a ballistic fingerprinting system is not restricted to those cartridge cases that can be identified by a trained firearm examiner.

FTI refers to a study made of 500 Glock semi-automatic pistols model 17 of caliber 9mm PARA. They obtain much higher correlation rates - in the 83% to 85% range - than the Smith & Wesson study conducted for the AB1717 Evaluation. This is probably due to the type of firearm. The Glock pistol leaves very characteristic and individual marks around the firing pin. They allow a firearms examiner to identify the weapon in which it was fired in a swift and straightforward way. This is confirmed in the AB1717 Evaluation [A, p.4-3]:"The Glock cartridge case, in particular, is generally easy to identify because of its strong breech face impression, firing pin aperture marks, firing pin shape, and striated detail." Because of the clear, detailed markings on cartridge cases, I use Glock cartridge case images to demonstrate firearms identification for laboratory visitors.

[12] J. De Kinder and M. Bonfanti, AFTE journal **31**, 318-323 (1999)

Performance test 1B - Correlation Item versus Database Size

A remark on the correlation times is given in Section 3.2. In conclusion "(…) correlation times are not a significant issue for a large database. (…) FTI appears to have scalable computer hardware that can accommodate large databases with minimal problems." [A, p.8-5]

Performance test 1C - Correlation Ranking Position as a Function of the Database Size

I fully agree with the analysis of the data as it was performed.

Performance test 2 - Cartridges Not in the Database

The IBIS-system presents its results as a list of candidate cartridge cases ordered by a correlation score. The score is only a relative value and is only significant for that one correlation. Its value for another correlation may be quite different. Hence, I do not know how to interpret the results from this test.

Performance test 3 - Ammunition Effect on Correlation

I fully agree with the analysis of the data as it was performed. "This test demonstrates the potential problems when different brands of ammunition are used and compared to that in the database. (…) It is not necessarily the fault of the correlation algorithm." [A, p.8-9]

Performance test 4 - Altered Breech Face

No remarks on the analysis of the (evident) results.

Performance test 5 - Correlation of 500 Sig Sauer pistols

This test was not performed at this time.

Performance test 6 - A Large Database query

This test was not performed at this time.

Performance test 7 - Breech Face Longevity Study

This test should be further elaborated and substantiated by a larger data set than just two firearms that fired up to 600 rounds of ammunition. The proposed test will provides quantitative data (ranking) on the longevity of firearms markings, not depending on the personal appreciation of a firearm examiner. A similar study has not yet been reported in literature.

Performance test 8 - Subclass Feature Effects on the Breech Face

This test was not performed at this time.

Discussion and Interpretation

The discussion goes on with the cartridges in a larger database of different firearms. The behavior of the position of the match is predicted. "It appears that the cartridge case has to rank in the 1st or 2nd place in the CHP database in order to be detectable in a much larger regional database of registered owners." [A, p.8-12] The general idea of this paragraph is correct.

(1) For a larger database filled with different firearms, the cut-off at the 2nd position is too severe and should be around 5. The corrected occurrence is 29.1%.

(2) For a larger database filled with firearms of the same model, the estimation is correct when the number of guns is substantially increased.

2. Remarks made in the Rebuttals from the Bureau of Alcohol, Tobacco and Firearms and from Forensic Technology Inc.

In this chapter, a number of relevant remarks from the BATF and from FTI which fall within the constraints of this study are discussed.

2.1 Common remarks

Subjective character of forensic firearms examination

Firearms-related marks are much more difficult to interpret and compare than DNA types or fingerprints. Unlike DNA, firearms-related marks from a single gun show some variation (see Section 2.2). All parties agree that the final identification of a firearm should be performed by a firearm examiner. A firearm examiner's conclusion will be based on his experience and is hence, subjective. The Association of Firearms and Toolmarks Examiners has set standards to provide a scientific support for a firearms examiner's conclusion. This is in contrast to e.g. DNA-analysis, which provides objective results and probabilities expressed in terms of population.

Rimfire cartridges, shotguns and rifles

Contrary to the AB1717 Evaluation, the IBIS-equipment can also accommodate rimfire cartridges and cartridges fired from shotguns and rifles. This remark was made in the rebuttal of FTI [C, p.7-8] and the BATF [G, p.14]. As a side-note:

- My laboratory does not deal with a substantial amount of rimfire cartridges to make an evaluation of performance of the IBIS equipment for these marks.
- If crimes are committed using revolvers, cartridge cases are missing most of the time. Hence, the importance to have them in the system is substantially reduced.
- In general, shotgun shells do not carry many marks, making them less suitable for a test project.
- AB1717 does not require any study or comments regarding long guns. (see FTI rebuttal [C, p.8]).

Altering firearms

The importance of altering the signatures of firearms, after a ballistic fingerprinting database is set up, is pure speculation. In my laboratory three cases out of a total of 1600 were recently seen, being about 0.2 %. Fig. 2 shows an example of the alteration of a firing pin. Two of them could still be identified back to a particular firearm using other marks present on the cartridge case. When a ballistic fingerprinting database will be operational, the number of occurrence will be higher than this value and is probably limited by the amount of weapons possessing erased serial numbers, which is about 10 % for my laboratory. I have no data for the firearms in California.

Whereas the BATF sees altering a firearm as a non-issue, it is a real problem: Any reduction in the potential of 'hits' such as caused by alteration to a firearm is of concern when evaluating the usefulness from a technical point of view of a 'gun sales database'.

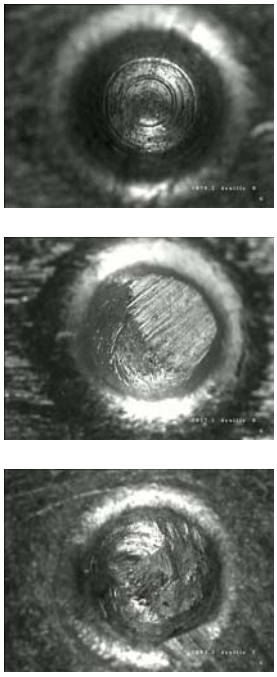

Fig 2 Example of the modifications on a firing pin. A positive identification between the cartridge cases could be made based on the breech face markings. The cartridge cases were found at a car jacking (1999), bank robbery (2000), and attempted murder (2001).

2.2 Rebuttal from the Bureau of Alcohol, Tobacco and Firearms

The BATF is the driving force behind the National Integrated Ballistic Information Network (NIBIN) program and provides networking of the existing IBIS-systems as well as purchase and deployment of additional systems and training of users in the United States. They wrote a rebuttal to the AB1717 Evaluation, mainly based on their acquired experience with NIBIN. This experience is solely based on the application of automated ballistic comparison systems to open case files. Their main remark is the choice of Federal ammunition for the test firing (already discussed in Section 2.2 of this paper).

The BATF rebuttal does not include a proposition how to go forward.

Two remarks still have to be addressed:

Correlation speed

The used computer array of 4 dual pentium stackable industrial computers used in this test is not the standard equipment for a stand alone IBIS system. Hence it is difficult to evaluate its performance compared to currently commercially available systems.

There seems to be an error in the AB1717 Evaluation, when comparing the results [A, p.1-4] for the Computer Capability and Speed and the experimental data [A, p.8-4 and p.8-5]. When extrapolating the data from the experiments to a hypothetical 100,000-cartridge case database, a correlation time of 50 minutes can be obtained and not 1.5 hours as printed on [A, p.1-4]. Given the current advances in computer technology, "it seems reasonable to

anticipate that computer processing will continue to get faster,…"[G, p.19]. In a different computer configuration, an SGI 2400 Computer equipped with a 15 co-processor unit would only need 20 minutes for this ([A, p.4-4 and p.4-5]).

I don't agree with the statement on [A, p.4-2] of the AB1717 Evaluation: "What is important to keep in mind is that while computer speed and computational power have made dramatic improvements in the last 10 years, the image algorithm is independent of this improvement. Its improvement is limited by the skill of the mathematician. Faster computers only enhance the calculation speed of the algorithm." While increasing the computer speed, the comparison algorithm will also be refined based on more experience in this field, the comparison algorithm will be updated to new mathematical techniques, new cameras for image capturing will be used,… One can even think of having 3D measurements made on the cartridge cases and bullets.[13] These improvements will increase the efficiency of the correlation.

Database uniformity and database size: effect on correlation

The BATF-rebuttal reads "The evaluation proceeds under the assumption that in a large database, actual hit exhibits will be pushed further down the correlation score list, as if other exhibits had better 'matching' detail than the actual 'hit'. This assumption is not supported by ATF examiner's experience in using IBIS. In actual fieldwork, IBIS correlation scores seem to actually improve with 'sister' test casings required, as the computer refines its search capability. Research listed at the end of this section describes this effect."[G, p.20]

This effect is typical when one is using neural networks. In actual fact, the possible use of neural networks in IBIS is glanced at in document [P], entitled "IBIS correlation results - improvements". This paper also reproduces a letter from FTI stating "Our engineering department does not possess a lot of experience in the use of neural network technologies. Your proposed use of the technology is interesting and will be considered in our plans to provide the IBIS-users with correlation results analysis tools." It is the experience of the IBIS-users from the firearms laboratory of my section that this feature is currently not yet implemented. So the remark of the BATF will only apply in the (perhaps near) future.

BATF's remark is also in contradiction to what is seen in the Evaluation AB1717 (see Fig. 8-8 on [A, p.8-6]) where the correlation position worsens with an increasing number of firearms in the database. The score does not improve with 'sister' cartridge cases present. The AB1717 Evaluation records that if a correlation score were at the 1st or 2nd position, it would stay there. If starting from a lower rank than the top two positions, its ranking would worsen with database size.

The use of a single quality control bullet [G, p.20], without any specification on its class characteristics does not present a statistical value for the quality of the system. It merely indicates that a number of components of the system are performing properly.

2.3 Rebuttal from Forensic Technology Inc.

For an introduction on Forensic Technology Inc., please see Section 2.3.

FTI denies the evaluation for a number of reasons:

[13] see e.g., *"Automated Comparisons of Bullet Striations Based on 3D Topology"* - J. De Kinder and M. Bonfanti, Forensic Science International **101,** 85-93 (1999).

1. An overly pessimistic and unsubstantiated view of the technology pervades the evaluation.
2. Criticisms of ballistic evidence should not equate to criticism of the technology.
3. The AB1717 Evaluation has critical omissions and errors and relies upon conjecture rather than evidence.

Forensic Technology Inc. has undertaken the effort to perform a similar study using 500 semi-automatic pistols GLOCK Model 17. The main results of this study are discussed in the Performance Test 1.

They conclude that : “[C, p.4]
1. Further study of this subject is necessary. FTI recommends that a pilot project be undertaken over the course of three years. Among the many things a pilot project could study, which the AB1717 Evaluation does not even consider is the filtering of large databases based on gun class characteristics, dates, regions and other factors to reduce sample sizes and increase processing efficiencies.
2. The developing IBIS technology will provide a near-term solution for the challenges posed by large databases.“

A number of remarks still have to be addressed:

Restricted use of the firearms investigated

"Aside from the fact the Evaluation made no attempt to determine what percentage or types of firearms would be included in this restricted group, this is an instance where FTI's technology has an advantage(…)" [G, p.16] It seems to me that this is more the work for FTI to investigate which firearms can be identified using their equipment.

Human comparisons

"The lack of comparison with human data is crucial and sufficient to raise doubts about the evaluation's main conclusions. The goal of IBIS is to reach the same conclusions that firearms examiners would have reached if they had the time to examine manually an entire database. Without comparing the computer results to the human results, the study is simply incomplete" [C, p.16]

I never heard that the fingerprint database or the DNA-database was verified through manual comparisons with the data. An additional study would not reach any supplementary result to what is already known in the scientific world. If however, FTI persists, the opportunity can be offered to them to finance such an additional study.

Time To Crime

The time to crime (TTC) is an important parameter that can be used to reduce the size of the ballistic fingerprinting database. If this parameter drops down in a relatively short time period, the firearms which were sold long before the average TTC have a negligible chance of being used in crime. In other words, it allows one to minimize the retention time of the data in the ballistic fingerprinting database without losing a substantial amount of its performance.

In the Crime Gun Trace Reports 2000 from the ATF, average TTC are mentioned per age of the offender and type of firearm [X, p.30-40]. The following results are obtained for semi-automatic pistols (4.5 years), revolvers (12.3 years), rifles (7.0 years), shotguns (7.6 years) and other firearms (7.1 years). The nationwide average TTC for all firearms for all ages of

offenders is 6.1 years. As this study averages over the whole U.S.A., regional differences can be expected. The study [X, p.32] particularly mentions Stockton, CA (9.2 years); San Jose, CA (9.0 years); Anaheim, Long Beach and Santa Ana (8.0 years) and Oakland, CA (8.0 years) as cities where the median time-to-crime is much longer than the overall city average. As all these cities lie within California, one can expect the average TTC for the State of California to be higher than the national value. More detailed data is required to determine correctly the median TTC for pistols in California.

Longevity of firearms markings

FTI reproaches the author(s) of the AB1717 Evaluation for not discussing the studies published by the AFTE on this subject [C, p.17]. However, one of the papers is cited in the footnote on [A, p.2-4] of the AB1717 Evaluation. It may be interesting to note that the referenced papers on this subject are reviewed in the article '*Influence of the Use of Firearms on their Characteristic Marks"* by Jan De Kinder and Monica Bonfanti[14], which contains general conclusions on this issue.

A study such as proposed by the AB1717 report provides quantitative data on the longevity of firearms markings, not depending on the personal appreciation of a firearm examiner. It is capable of showing a possible drift of the correlation results after a number of cartridges have been fired by the firearm.

Over-representation of certain firearms

The AB1717 Evaluation concluded that there is an increased potential as the database increases in size, for a firearm to be over-represented in that database, making it more difficult to correlate matches. I agree with the remark from FTI [C, p.18] that the current study does not provide any support for this statement. This point was clarified by Fred Tulleners lateron.[15]

Larger databases

FTI objects [C, p.19] to "As progressively larger numbers of similarly produced firearms are entered into the database, images with similar signatures should be expected that would make it more difficult to find a link. Therefore, this increase in database size does not necessarily translate to more hits". This is corroborated by Figure 8-8 [A, p.8-6] of the evaluation. This figure shows that when increasing the database size, the ranking of a cartridge case decreases substantially. The current analysis protocol (see section 2-4 of this review) has to be adopted in order to still be able to detect them as a hit.

It must be stressed however, that FTI mostly develops its system using small sized databases. I refer to their presentation during the AFTE 2002 meeting in San Antonio. A test set of 150 firearms was used and 'statistically extrapolated' to 5000 firearms. This technique can at least be called 'unusual', especially for a company that should have access to large quantities of firearms data.

[14] AFTE journal **31**, 318-323 (1999)

[15] If different firearms are manufactured using similar methodology, they will end up possessing the same class characteristics. Hypothetical, 1,500 Smith & Wesson pistols could look like 1,500, resulting in a database of 3,000 firearms which have the same class characteristics. (Information received by email on November 28, 2002).

Further Study

1. The results were obtained for one particular type of semi-automatic pistol. Much better correlation results were obtained by FTI for Glock pistols. It would be good to study a third firearm to improve our view on the possibility of ballistic fingerprinting. Also performance test nrs 5, 6, 7, and 8 of the AB1717 Evaluation still have to performed.

2. Different protocols should be developed and tested to optimize the correlation performance of the ballistic fingerprinting. This can be done by varying (at least) two parameters:
   - Via an optimization of the choice of ammunition used to test-fire the guns. By correlating different types of ammunition fired through a single firearm, the best suitable ammunition for test firing can be determined.
   - More than one cartridge cases can be entered per firearm. This allows to account of a certain variation in the marks left on different makes of cartridge cases. In a previous study, this protocol increased the correlation performance strongly.

3. The technology used can still be improved. This task lies with FTI as the manufacturer and the owner of the intellectual property behind the IBIS-system. They have several directions they can take:
   - use other marks on the cartridge cases to assist in the correlation procedure,
   - improve the technology of the system (camera, correlation algorithm,..),
   - study the possibility of recording 3D images.

4. Evaluate the possibility of having a serial number / bar code imprint on the breech face of a firearm, which is transferred to the fired cartridge. Is this practically possible ? Could this be imposed by law as part of SB15 ?

5. Create a test project, similar to the demands by FTI, for a single caliber of ammunition and for a couple of years. Evaluate the performance of the system as well as its consequences for the police inquiries. This subject can also be studied by following up the results from the ballistic fingerprinting programs in the States of Maryland and/or New York. No results of their project have been presented so far at conferences or published in scientific journals.

6. Determine a correct median Time To Crime (TTC) for California. Determine the number of firearms in all calibers per year. This will allow one to determine the maximum size of similar firearms present in the system.

Acknowledgements

I would like to thank the members of the Ballistics Section of the NICC/INCC for providing the supporting data and passing the experiences with the IBIS-system, especially Hugues Thiebaut.