Autopsy of a Crime Lab

Exposing the Flaws in Forensics

Brandon L. Garrett
A Report

CITY OF PHILADELPHIA
Police Department
Office of Forensic Science
843-449 N. 8th Street
Philadelphia, PA 19123
(215) 685-3101
Firearms Identification Unit Laboratory Report

Case Information:
Investigation type:
Location:

Evidence Inventory:

<table>
<thead>
<tr>
<th>Item</th>
<th>Property Receipt (PR)</th>
<th>Quantity</th>
<th>Evidence Description</th>
<th>Location</th>
</tr>
</thead>
<tbody>
<tr>
<td>Q1</td>
<td></td>
<td>1</td>
<td>Fired Cartridge Case</td>
<td></td>
</tr>
</tbody>
</table>

Methods of Analysis
A visual and physical examination was conducted on all submitted evidence. A microscopic comparison was conducted on all suitable ballistic evidence (unsuitable items are noted in analysis results).

Results/Conclusion:
This is a supplemental report for report dated 2017.
Item Q1 One (1) Fired cartridge case FCC-4. Winchester brand (nickel), 9mm Luger, displaying a Glock type firing pin impression.
Conclusion:
Microscopic examination was conducted with the following results.
Fired cartridge case FCC-4 was identified as having been fired in P-1.

Remarks:
NIBIN database search results: Not applicable
Laboratory fee: $150.00
Identifications are made when there is an agreement of all discernible class characteristics and sufficient agreement of individual characteristics.

Inconclusive results are reported when items could not be identified or eliminated due to an absence, insufficiency, or lack of reproducibility of individual characteristics.

Eliminations are made when there is a disagreement of discernible class characteristics and/or individual characteristics.

Disposition of evidence: Evidence shall be maintained by the PPD in accordance with all applicable directives, standards and/or legal requirements.

**QA/QC Statement:**

This report accurately reflects the findings and opinions of the examiner who performed the analysis. All of the examinations and tests were performed in compliance with validated and industry-approved procedures and standards.

Ronald Weitman #... Examiner
Police Officer

Robert Stott #... Co-Examiner
Police Officer
What was the conclusion?

A sufficient correspondence of individual characteristics will lead the examiner to the conclusion that both items (evidence and tests) are from the same source.

Fired cartridge case FCC-4 was identified as having been fired in P-1.
Data webpage
http://www.convictingtheinnocent.com
Types of Forensic Testimony

- Voice Spect. - 100% (1/1)
- Bite mark - 71% (5/7)
- Shoe print - 16% (1/6)
- DNA - 17% (3/18)
- Fingerprint - 5% (1/20)
- Hair - 39% (29/75)
- Serology - 58% (67/116)
Who did the work? How proficient were they? Where they qualified?
The Impact of Proficiency Testing
Information and Error Aversion

- Gregory Mitchell & Brandon Garrett

We commissioned Qualtrics to recruit nationally representative sample with respect to gender, race/ethnicity, age, income, and geographic region in the United States. Total of 1,450 adults participated in the study, which took less than 15 minutes. In addition to asking demographic questions, we gave an objective numeracy test to participants.

The description of the case was kept simple to keep the Participants focused on the fingerprint evidence itself. The survey software assigned participants to one of 14 conditions with five proficiency levels and three error types, as well as a control in which the examiner received a perfect score on proficiency (with no errors) and a control condition with no proficiency information provided.
The Impact of Proficiency Testing
Information and Error Aversion
Fired cartridge case FCC-4 was identified as having been fired in P-1.
How was the conclusion reached?

<table>
<thead>
<tr>
<th>Item</th>
<th>Property Receipt (PR)</th>
<th>Quantity</th>
<th>Evidence Description</th>
<th>Location</th>
</tr>
</thead>
<tbody>
<tr>
<td>Q1</td>
<td></td>
<td>1</td>
<td>Fired Cartridge Case</td>
<td></td>
</tr>
</tbody>
</table>

Identifications are made when there is an agreement of all discernible class characteristics and sufficient agreement of individual characteristics.
What methods were used?

Methods of Analysis

A visual and physical examination was conducted on all submitted evidence. A microscopic comparison was conducted on all suitable ballistic evidence (unsuitable items are noted in analysis results).
Firearm-Related Toolmark (FATM) Analysis

Generally:
• Markings on bullets and cartridge cases believed to be associated with a crime are compared against . . .
• Markings on bullets and cartridge cases test fired from a firearm believed to be associated with a crime . . .
• Do they match? Or, “sufficiently” match?
• Evidentiary “show up.”
Terms

- Scratches ("striations" or "striae") and impressions left by the interior surface of the firearm on the bullet or cartridge case.
- Lands and Grooves
Firearm examiners recognize three categories of markings:

- Class;
- Subclass;
- Individual.

**Individual**: Characteristics that firearms examiners believe are unique to a firearm, resulting from some combination of irregularities in the firearm machining process and imperfections that emerge during the subsequent use of a firearm.
Firing pin and breech impressions
Comparison (AFTE publications)
Changes in manufacturing decrease “individual” and increase subclass marks

• “[M]ass production of guns has replaced hand-manufacturing.”  
  US v. Mouzone

• Manufacture under “precisely controlled” conditions imparts “recurring patterns” of marks.  D. Baldwin, Statistical Tools

• Tools used to create firearms have become more durable, enabling their use in ever-larger production runs.  P. Kirk, Crime Investigation
“Theory of Identification”

• Association of Firearms and Toolmark Examiners (AFTE) instructs practitioners to use the phrase “source identification” to explain what it means when they identify “sufficient agreement” when examining firearms (AFTE, 1998).

“Agreement is sufficient when it exceeds the best agreement demonstrated between toolmarks known to have been produced by different tools and is consistent with agreement demonstrated by toolmarks known to have been produced by the same tool.”

If match meets this sufficiency “standard”, the likelihood of a coincidental match is “so remote as to be considered a practical impossibility.”
• 2008 NRC committee, responding to NIJ’s request, found “the validity of the fundamental assumptions of uniqueness and reproducibility of firearms-related toolmarks” had not yet been demonstrated.

• 2009 National Academy of Sciences Report - categorical conclusions regarding firearms or toolmarks were not supported by research, and that instead, more cautious comparative claims should be made.

• The report stated the “scientific knowledge base for tool mark and firearms analysis is fairly limited (NAS 2009, pg. 155).”

• An examiner makes “a subjective decision based on unarticulated standards and no statistical foundation for estimation of error rates (NAS 2009, pg. 153-54).”

• The AFTE theory of identification is inadequate and does not explain how an expert can reach a given level of confidence in a conclusion.
PCAST Report (2016):

• Validity as applied would also require, from a scientific standpoint, that an expert testifying on firearms analysis (1) has undergone **rigorous proficiency testing** on a large number of test problems to measure his or her accuracy and discloses the results of the proficiency testing and (2) discloses whether, when performing the examination, he or she was aware of any other facts of the case that might influence the conclusion.
PCAST Report (2016):

• Only one proper study has been conducted.

• We identified one notable advance since 2009: the completion of the first appropriately designed black-box study of firearms. The work was commissioned and funded by the Defense Department’s Forensic Science Center and was conducted by an independent testing lab (the Ames Laboratory, a Department of Energy national laboratory affiliated with Iowa State University).

• The false-positive rate was estimated at 1 in 66, with a confidence bound indicating that the rate could be as high as 1 in 46. While the study is available as a report to the Federal government, it has not been published in a scientific journal.

• The inconclusive rate was 33.7 percent

• Other studies were “closed-set” designs – where the correct answer is always present in the collection.
<table>
<thead>
<tr>
<th>Latent</th>
<th>4P1</th>
<th>Known:</th>
<th>Johnny Uber</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>Case #:</td>
<td>2017FEL1234</td>
</tr>
</tbody>
</table>

Verified: MP
What Conclusions were Reached?

Q And you look for points of identification that match up?

A Yes.

Q And how many total points of identification do you need to make a positive identification?

A Eight.

Stephan Cowans’ Trial

A That they were identical.

Q Whose print was it?

A Stephan Cowans’.
State v. Williams (COA 2018)

• “[T]he four evidence cartridge cases were fired from the same firearm as the test firings that I found. So they were fired from the same gun that I test fired.”

• COA:
  • [W]hile Mr. Bishop did not qualify his opinion with “to a reasonable degree of certainty,” he also never uttered the words “unique as to each gun that’s made” or “exclusive identification” two phrases defendant refers to extensively in his brief as the alleged claims of certainty that amounted to false overstatements of reliability. In fact, it was defense counsel, not Mr. Bishop, who chose to use the exact phrases defendant challenges on appeal.
  • At no time, either on direct or cross-examination, did defense counsel object to any portion of Mr. Bishop’s testimony or dispute the reliability of his expert opinion.
“Source identification’ is an examiner’s conclusion that two toolmarks originated from the same source. This conclusion is an examiner’s decision that all observed class characteristics are in agreement and the quality and quantity of corresponding individual characteristics is such that the examiner would not expect to find that same combination of individual characteristics repeated in another source and has found insufficient disagreement of individual characteristics to conclude they originated from different sources. (p. 2)
PCAST Report (2016):

- Only one proper study has been conducted.
- We identified one notable advance since 2009: the completion of the first appropriately designed black-box study of firearms. The work was commissioned and funded by the Defense Department’s Forensic Science Center and was conducted by an independent testing lab (the Ames Laboratory, a Department of Energy national laboratory affiliated with Iowa State University).
- The false-positive rate was estimated at 1 in 66, with a confidence bound indicating that the rate could be as high as 1 in 46. While the study is available as a report to the Federal government, it has not been published in a scientific journal.
- The inconclusive rate was 33.7 percent
- Other studies were “closed-set” designs – where the correct answer is always present in the collection.
The scientific criteria for foundational validity require that there be more than one such study, to demonstrate reproducibility, and that studies should ideally be published in the peer-reviewed scientific literature. Accordingly, the current evidence still falls short of the scientific criteria for foundational validity.

Whether firearms analysis should be deemed admissible based on current evidence is a decision that belongs to the courts. If firearms analysis is allowed in court, the scientific criteria for validity as applied should be understood to require clearly reporting the error rates seen in the one appropriately designed black-box study. Claims of higher accuracy are not scientifically justified at present.
How reliable is the method?

- 1443 conclusive examinations
- 22 false positives
- False Positive rate of 1 in 66 cases.
  → Upper bound of 1 in 46 cases (2.2%)
  → Available at https://www.ncjrs.gov/pdffiles1/nij/249874.pdf
- But also subjectivity issues:
  - 218 examiners
  - Each given 10 different source sets
  - 44% - all 10 sets suitable for definitive opinion
  - 20%- all 10 sets are inconclusive
  - 35%- some inconclusive, and some suitable
How about Bias?

• It is important to note that, for a verification program to be truly blind and thereby avoid cognitive bias, examiners cannot only verify individualizations. As the authors of the FBI black-box study propose, “this can be ensured by performing verifications on a mix of conclusion types, not merely individualizations”—that is, a mix that ensures that verifiers cannot make inferences about the conclusions being verified.

• We are not aware of any blind verification programs that currently follow this practice.
How about Judges? Judicial Instructions?

- Although “commissions come and go” there is “mounting judicial, and public concern...”
Federal judges have required that examiners opine only that it is “more likely than not” that the ammunition could have come from the defendant’s firearm (Glynn, 2008).

Others have required that the examiner limit conclusions to a “reasonable degree of ballistic certainty” (Montiero, 2006, Diaz, 2007).

Another judge limited testimony to observing that the markings were “consistent” (Willock, 2010).

Firearm “could not be eliminated” (Goodwin-Bey, 2016)

Most recently, a judge has ruled that a firearms expert can say no more than the firearm “cannot be excluded” (Tibbs, 2019).
How do jurors evaluate such information?
Figure 1. Proportion of guilty verdicts (with 95% confidence intervals) in each experimental condition.
Work in Progress: Firearms Testimony and Jurors’ Evaluation of Forensic Evidence

Brandon Garrett, Nicholas Scurich & William Crozier

In a pre-registered experiment with 1,400 mock jurors (n=200 per cell), selected to be jury eligible and census representative by age, race, gender and geographic region, we presented participants with a case containing firearms evidence, and varied the wording of a firearm examiner’s conclusion. Participants rated the evidence, expert, and rendered a verdict.

Variation in conclusion language did not affect guilty verdicts. In contrast, a more cautious conclusion that an examiner “cannot exclude the defendant’s gun,” did lower verdicts.
Prior work: *How Jurors Evaluate Fingerprint Evidence: The Relative Importance of Match Language, Method Information and Error Acknowledgement*

Brandon Garrett & Gregory Mitchell


Fingerprint evidence benefits from common beliefs and background assumptions in uniqueness and reliability of fingerprint identification.

Language may not be important once jury is told a match was made (Match = Match to Exclusion of All Others = 100% Certain = Other Source a Practical Impossibility...)

Error statements by forensic experts should be given greater attention by courts and researchers.
Thank you

• bgarrett@law.duke.edu