Validation Study of Electrochemical Rifling

By: Charles S. DeFrance and Michael D. Van Arsdale, Federal Bureau of Investigation, Washington, D.C.

Key Words: electrochemical machining, electrochemical rifling, ECR, validation study

ABSTRACT

Electrochemical machining is being used by Smith & Wesson to rifle many of its revolver barrels. This paper provides a description of this manufacturing process and a study that was conducted to evaluate whether or not these barrels will mark bullets in a repeatable and unique manner. This validation study of firearms/toolmarks identification as it applies to electrochemical rifling found that this manufacturing technique does produce unique, reproducible, and identifiable microscopic marks.

INTRODUCTION

Since 1993, Smith & Wesson has been using an electrochemical machining technique to rifle most of their revolver barrels. The only revolver barrels that are still broach rifled are .22 caliber barrels and ported barrels. The manufacture of electrochemically rifled (ECR) barrels begins with the same steps as conventional broach rifling. The barrels are drop forged from bar stock, annealed, and wheel abraded to remove scale. During the annealing process the barrels have a tendency to bend and are therefore put through a straightening operation. The barrels are next drilled and reamed using conventional machining tools and the forcing cone is made with a tapered reamer. The barrels are then ready for rifling.

The electrochemical rifling machines are made by Surftran and were specifically designed for Smith & Wesson. Each machine runs two independent workstations, each one with a single electrode manufactured by Mechanical Plastics. They are constructed of a two-inch long plastic cylinder with metal strips spiraling down its exterior. The metal strips are in the desired dimensions of the grooves, are at the appropriate rate of twist (1 turn in 18.75 inches for .357 Magnum), and are slightly inset in the plastic cylinder. The barrel is placed in the machine and is held stationary. The electrode is placed into the barrel and both are submerged in an electrolyte (sodium nitrate). The electrode travels down the barrel and rotates at the desired rate of twist. As current passes from the negatively charged electrode (cathode) to the positively charged barrel (anode), the metal is removed by electrolysis to produce the grooves by duplicating the shape of the electrode. During this operation the electrolyte flows through the barrel under pressure to remove the reaction products. This prevents the build up of reaction products on the electrode. Because the metal strips on the electrode never come in physical contact with the barrel and reaction products are not given the opportunity to build up, the electrode does not require any cleaning or maintenance. In fact, electrodes are only retired when the plastic core, which contacts the barrel to provide proper spacing and centering, wears over time. An electrode will usually remain within the tolerance of 2 thousandths of an inch concentricity for approximately 3000 inches of barrel. During our tour of the Smith & Wesson factory, they were rifling six-inch .357 Magnum caliber barrels and the ECR process took about 60 seconds per barrel.

While touring the facility, Smith & Wesson generously provided five consecutively rifled six-inch .357 Magnum caliber barrels. These barrels were rifled in the presence of one author. Each barrel was numbered in order of production, wrapped to avoid damage during transport and taken to the laboratory for further examination and testing.

PROCEDURE

The five consecutively rifled barrels were numbered in the order of manufacture. Each barrel was test fired on the same Smith & Wesson revolver, a Model 681. However, the marks present on these first sets of bullets were difficult or impossible to identify. It is believed that this is due to rapid wear of the new barrels before the microscopic characteristics stabilize. This phenomenon has been previously documented in new, unused barrels in studies conducted by Murdock[1] and Matty[2]. Their studies required a couple sets of test fires before the marks began to stabilize. However, the marks in the ECR barrels did not seem to be stabilizing as quickly. To avoid any possibility that changing marks might interfere with the study, fifty rounds of jacketed ammunition were fired from each barrel to represent the "break-in" period.

After the break-in period, test samples were fired and collected from each barrel. Microscopic comparisons showed that the barrels were reproducing their microscopic characteristics on the test fires. These samples were .357 Magnum caliber, 158 grain jacketed soft point bullets. For each barrel, six test bullets were collected. The fired bullets were randomly lettered and placed into envelopes marked with the respective barrel number.

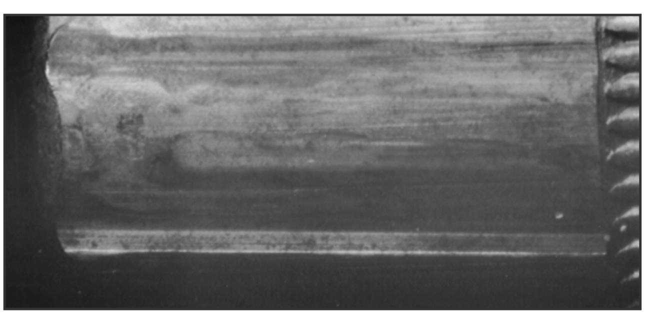
Three different tests to be conducted by a qualified examiner were created from the test fired specimens. Each test consisted of five bullets that were randomly selected from the envelopes such that one bullet from each barrel was represented. Two additional bullets were added to each test. The additional bullets provided for at least two possible identifications. However, in one test the two additional bullets had both been fired from the same barrel and therefore three identifications were possible.

Each test was given to a qualified Firearms-Toolmarks Examiner in the FBI Laboratory. These tests required twenty-one different bullet to bullet comparisons. Each examiner was asked to fill out an answer sheet and mark each comparison they made as an identification, no conclusion, or exclusion. For every identification, they were to provide information as to whether their identification was based on marks present in the land impressions alone, the groove impressions alone, or both lands and grooves independently. This distinction was made because marks produced by the lands are a result of the reaming process, while the grooves produce marks that are a result of the ECR process. Thus, different areas on the bullet represent different manufacturing processes. A total of nine examiners completed a test.

OBSERVATIONS

When the barrels were examined in the laboratory it was noted that the rifling had the general appearance of conventional rifling. However, upon closer inspection it was noticed that the shoulders between lands and grooves were not as sharp as commonly seen in broached, button, or hammer forged rifling. This was also apparent upon examination of the test fired bullets, which also had a less defined shoulder between land and groove impressions (Figures 1 through 3).

Figure 1



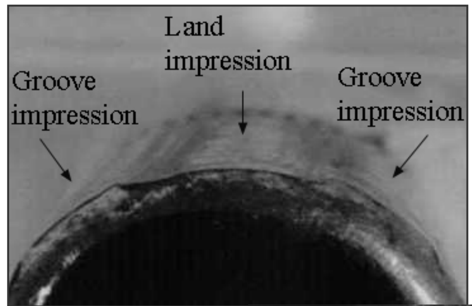
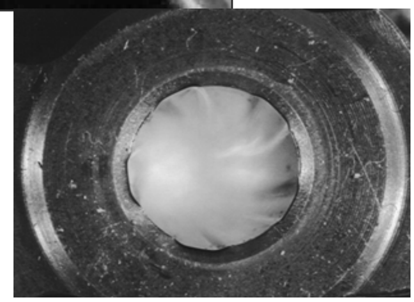


Figure 2

Figure 3



Figures 1 & 2. Photographs of a test fired bullet showing the rounded shoulders of the land and groove impressions. Figure 2 is the base of the bullet. Figure 3 is a photograph of the muzzle end of a test barrel, showing the rifling.

The general rifling characteristics of these bullets were measured and are listed below:

Five Grooves, Right Twist
Land Impression Width: 0.097”-0.100”
Groove Impression Width: 0.116”-0.120”

RESULTS

With the exception of one of the authors, all nine of the qualified Firearms-Toolmarks Examiners in the FBI Laboratory participated in this study. Upon completion of the tests, the results were collected and analyzed. The responses from the nine examiners included no false identifications or false eliminations. All examiners reported that the identifications that they made could be made independently on the land or groove impressions. In three of the tests there was a true identification that was marked as a “no conclusion.” However, only false positive or false negative responses were considered incorrect since a “no conclusion” does not exclude the possibility that the bullets could have been fired from the same barrel.

CONCLUSIONS

Based upon one author’s personal observations during the comparison of test fired bullets from each barrel, it was clear that marks were consistently reproduced. Further, these reproduced marks were clear on both land and groove impressions, which is important since they would each be the result of two different manufacturing processes (Figures 4&5).

The results of the tests are also very positive. Without exception, all the examiners reported correct results. There were no false identifications reported which clearly indicates that the marks left on the bullets are unique to a specific barrel. This was expected in reference to marks produced by the lands, which are the result of a reaming operation. Hall[3] has previously documented the uniqueness of marks produced by reaming. The results in this study serve to further support those reported in previous studies.

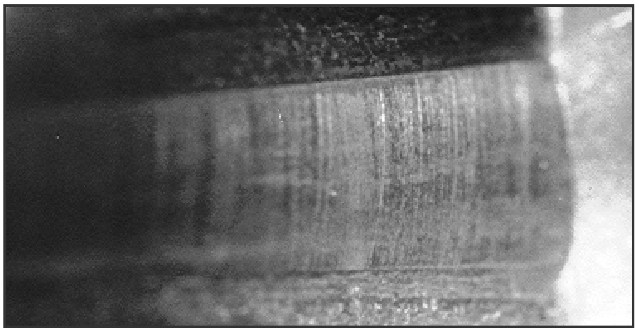
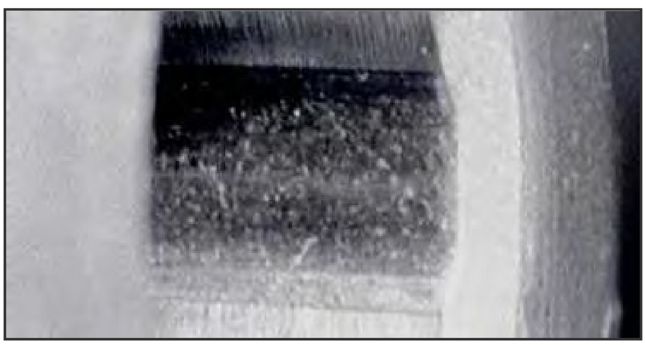


Figure 4. The marks left by the reaming process are visible on the top of the land in a test barrel.

Figure 5. The marks left on the grooves are visible



here. A speckled pattern is visible from the removal of metal during the ECR process.

Additionally, each examiner reported that it was possible to effect identifications based on the marks in the groove impressions alone. These marks are the result of the electrochemical rifling. This clearly indicates that the electrochemical rifling does produce unique and identifiable microscopic marks.

ACKNOWLEDGMENTS
The authors would like to thank Smith & Wesson for providing the barrels for testing and for placing their extremely knowledgeable staff at our disposal.

The authors would also like to thank all the examiners in the Firearms-Toolmarks Unit who took the time from their busy caseloads to assist in this study.

REFERENCES

1) Murdock, John E., "A General Discussion of Gun Barrel Individuality and an Empirical Assessment of the Individuality of Consecutively Button Rifled .22 Caliber Rifle Barrels," ATFE Journal, Volume 13, No. 2, 1981, pp. 84-95.
2) Matty, William, "A Comparison of Three Individual Barrels Produced from One Button Rifled Barrel Blank," AFTE Journal, Volume 17, No. 3, July 1985, pp. 64-69.
3) Hall, E. "Bullet Markings From Consecutively Rifled Shilen DGA Barrels," AFTE Journal, Volume 15, No. 1 January 1983, pp. 33-53.

Empirically Determined Frequency of Error in Cartridge Case Examinations Using a Declared Double-Blind Format

By: Angela Stroman, Criminalist, California Department of Justice, Bureau of Forensic Services, Redding Regional Crime Laboratory

Keywords: bias, breech face marks, cartridge case, blind testing, declared testing, double-blind, empirical study, error, manufacturing processes, null hypothesis, proficiency tests, scientific method, reliability, single-blind, Smith & Wesson 4006TSW pistol, subclass characteristics, validation study

ABSTRACT

Although the discipline of firearm and tool mark identification has been accepted by courts since the early 20th Century, it has come under serious criticism from scientific and legal experts over the past several years. One of these criticisms is that the forensic science community, and the field of firearm and tool mark examination in particular, has not done enough to control bias in empirical research, such as participating in blind testing. Numerous examples of firearm and tool mark validation studies can be found in the relevant scientific literature. Many of these are single-blind studies involving the use of consecutively manufactured tools. A literature search of the AFTE Journal archives revealed only two studies that involved declared double-blind testing of firearm and tool mark examiners. The goal of this study was to empirically determine the frequency of error in firearms identification results produced by qualified examiners who were presented with realistic samples using a declared double-blind testing format.

Introduction

The field of firearm and tool mark identification has a long history that is grounded in the scientific method [1, 2]. In practice, forensic science is an applied science and may appear to lack the hallmarks of the scientific research on which it is based. Several academic researchers (non-firearm examiners) have alleged that even while firearm and tool mark examiners continue to perform research in their field, they are not doing so in a scientific manner. Several critiques about the pattern and impression disciplines within forensic science have been published following the release of the 2009 National Academy of Sciences (NAS) report on strengthening the forensic sciences. Many of these articles express the criticism that the empirical studies that have been published are: 1) not blind or double-blind; and 2) unreliable [3, 4, 5, 6, 7].

This article is not intended to provide a comprehensive literature review of how or why firearm and tool mark identification should be considered a science in light of its recent criticisms. It will, however, explain what blind studies are and their significance in minimizing sources of bias in research, the application of the scientific method to firearm and tool mark identification, and present the findings of the author's research, which incorporated these concepts.

Declared vs. Blind Testing

Blind studies are often used in scientific fields in order to obtain information about a test, product, or a theory with as little bias, conscious or subconscious, as possible. The basis of a blind study is that the people taking the test are unaware of certain information that might lead them to be biased toward a particular response [8].

The easiest way to describe a blind study is to provide an example: a manufacturer of peanut butter asks a subject to taste test two different brands of peanut butters and rate them. The peanut butters are presented to the tester in plain jars, labeled "A" or "B", eliminating the possibility of bias due to brand preference.

However, there is not a clear consensus in the scientific literature as to what constitutes a "blind" test. Two noted critics of the manner in which testing is conducted in the forensic sciences, Saks and Koehler, acknowledge that their use of the phrase "blind examination procedures" can have a variety of meanings [9]. For instance, in the first part of a two-part survey on the feasibility of external blind DNA proficiency testing published in 2003, Peterson, et al. describe a test as being "declared" (or "open") if the subjects know they are being tested, and a "blind" test as one in which the respondents are not aware they are being tested [10]. The 2009 NAS report on forensic science echoes these definitions [11]. The website of the National Institutes of Health (NIH), on the other hand, describes two types of blind clinical trials: 1) "single-blind" studies, in which only the patients are not told what drug they are being given; and 2) "double-blind" studies, in which the patient as well as nearly all of the hospital staff involved in administering the test do not know this information [12]. Yet, in both of these last two situations, the patient still knows they are participating in a test. Throughout

Date Received: October 11, 2011
Peer Review Completed: December 20, 2011

the remainder of this paper, the terms "single-blind" and "double-blind" will be used in the same general context as the NIH descriptions unless otherwise specified. To use a clinical trial as an example, a patient agrees to participate in a test of a new pharmaceutical drug knowing that they will either receive the actual drug or a placebo, but not which one they actually do receive. Because the test is blind in this respect, the doctor is able to evaluate the efficacy of the drug based on the progress of those patients in the group that received the actual drug versus the group which received the placebo (the control group), without concern for any cognitive bias that may have been introduced had the patients been aware of which type of dose they had been given. In a double-blind scenario, the doctors would also not know which patients were in which group until the end of the study. In this model, the fact that the patients know they are participating in a study does not diminish the blind aspect of the test, since the crucial factor is that both the patients and doctors do not know to which treatment group a particular patient is assigned.

Similar to this example, when a forensic examiner participates in a research study or a declared proficiency test, they are aware that they are being tested. The participants are not aware of what the true answers are or how the tests were created, and therefore a certain amount of confirmation bias has been removed. However, single-blind studies still allow the test giver to potentially interject bias either intentionally or unintentionally. For example, a test is made by a supervisor in the forensic lab and then given to an examiner. The examiner reports their results back to the supervisor, who then displays some unintentional visible or verbal queue to the examiner when the reported results were not what the supervisor anticipated; the examiner is now biased as to what the results might be should the examiner be allowed to revise them before officially submitting them. Alternatively, the examiner could present the findings to the supervisor, who then intentionally tells the examiner, "You may want to go back and look at Item 1," indicating to the examiner that their initial answer was wrong and they need to re-evaluate.

Most proficiencies and research projects consist of single-blind testing. These tests can limit bias if the administrators do not work in the same office as the participants. In the case of externally-administered proficiency tests, such as those provided by Collaborative Testing Services (CTS), the test creators are specifically tasked with making and distributing the tests and never have any, or at least very limited, interaction with the participants.

In a single-blind test, the fact that the participant knows they are being tested can also play a significant role in their responses to the test. The examiner may feel more pressure to perform accurately and get the correct answer, which may lead the examiner to treat the test in a different manner than real casework. For example, when an examiner knows they are taking a CTS proficiency test, they may try harder to get the right result because they know there is a lot of personal accountability riding on it. Furthermore, in addition to the possible stigma of getting the wrong answer, the examiner also knows that the laboratory will likely use the results of these tests as a standard for how well the examiner is performing their examinations, as often dictated under the laboratory's quality assurance guidelines. This can be problematic, since CTS considers a "correct" answer to be one that is among the consensus of the responses they receive from participants and not necessarily what the true answer is based on how the test was prepared (which is known by CTS) or the quality of a particular test sample [13]. Therefore, if an examiner submits an answer (e.g. an inconclusive result) that does not meet the consensus, regardless of the reason, the laboratory can view the conflicting result as a quality issue and require the examiner to be retrained. Prosecutors and defense attorneys may also inquire about an examiner's past CTS results as a matter of practice before or during the examiner's testimony. If the examiner has any results that are different than the consensus, they may need to explain the inconsistency every time they testify. Potential repercussions such as these may lead examiners to treat declared proficiencies differently than normal casework.

While CTS proficiencies can be considered single-blind tests, they typically do not accurately reflect the full range of difficulty experienced in real-life firearm and tool mark casework, because they consist of actual fired bullets and cartridge cases or other non-firearm tool marks presented in pristine condition for the sake of reproducibility. Damaged test samples or those with borderline microscopic agreement are inherently very difficult to mass produce consistently using live ammunition. CTS readily admits the limitations of their tests, and has accordingly issued a statement discouraging the use of its tests to calculate error rates for the field [14]. However, in the absence of more controlled studies, proficiency test results such as those from CTS, when properly considered in the context of their limitations, currently appear to provide the best source of data from which to project meaningful false-positive error rates for the field for use in defending against court admissibility challenges [15]. Various commentators have debated the merits of using proficiency tests in this manner; these viewpoints have been summarized by Kaye, et al [16]. Still, since it is recognized that declared proficiency tests are neither designed nor very suitable for determining error rates for individual examiners or a profession as a whole, it has been proposed by at least one study that data be considered from several sources, one of which being realistic, blind proficiency testing [17].

A review of articles published in the AFTE Journal indicates that most of the past research and validation studies performed in regards to firearm and tool mark identification can be considered to be, at the least, single-blind [18, 19, 20, 21]. Despite the potential pitfalls for bias described above, single-blind studies can still give valid scientific results if conducted with those pitfalls in mind. However, an ideal approach to addressing concerns of bias in validation studies would appear to be the double-blind model. True double-blind (or simply "blind") studies are those in which the test administrators do not know which participants received which samples and the participants are unaware they are being tested. Even though this type of testing has the greatest potential for minimizing sources of bias, it is difficult and costly to implement. To perform such a test in a forensic laboratory would involve submitting the test samples to the laboratory as if they were routine evidence samples, to include misrepresenting who was actually submitting the test so no one in the lab administration, let alone the examiner, would be aware of the test. In this scenario, the laboratory administrators would have contracted with the external test provider prior to the test, but would not know anything about the test or even be aware of when it was being worked on in the laboratory. This type of testing could be implemented in a variety of ways, with either just the target examiner or the entire laboratory being kept unaware of the test. One major advantage that has been claimed of true double-blind testing is that it can be used to test "the whole system": the processes a laboratory uses from the time it receives the "evidence" until the time the results are reported out. Obviously, the feasibility of implementing a truly double-blind system of proficiency testing on any significant scale is low, as it would need to involve extremely careful test preparation and the participation of local law enforcement agencies to successfully "deceive" the target laboratory and examiner without raising any suspicions that it might be a test. Due to the intricacies of administering such a test, the cost per test would be prohibitively expensive for any but the most well-funded laboratories. For instance, it has been conservatively estimated that the cost to implement a large-scale (150 laboratories), national double-blind proficiency testing program for forensic DNA analysis would run approximately $1,400 to $10,000 per test, depending on the test provider and the variables of the testing model used [22].

A more practical compromise to true double-blind testing is to incorporate as many of the elements of double-blind (i.e. "blind") testing into a "declared" test, one in which the participants know they are participants. A search of the AFTE Journal archives revealed two studies that deal with declared double-blind testing of individuals: a 2003 validation study performed by Bunch and Murphy, and a 2009 validation study by Giroux [23, 24].

The design of the Bunch and Murphy study, which measured examiners' ability to identify cartridge cases from consecutively manufactured Glock slides, incorporated four principles: 1) return of the test was mandatory; 2) the tests could not be traced back to individual examiners, so the participants would not "try harder" on their test samples for fear of reprisals should they commit any errors; 3) the tests were all different, eliminating the possibility that the tested examiners could reveal or obtain the correct answers by discussing the test amongst themselves; and 4) the test was called "double-blind" by the authors in the respect that the participants were unaware of the correct answers, and that the test administrators (Bunch and Murphy) did not know who received which test specimens. As a further bit of caution to reduce bias, the test answers were reportedly kept under lock and key until the test was completed. Unlike a true double-blind test, but similar to the majority of validation studies within the field of firearms identification, the participants in the study were aware of the fact that they were being tested and therefore any potential bias created by this knowledge was still present [25].

In Giroux's study, the ability of examiners to identify tool marks made by consecutively manufactured screwdrivers was investigated. Giroux followed the experimental design set forth by Bunch and Murphy insofar as: 1) the return of the test was mandatory; 2) the answers to the test were not traceable back to an individual examiner; 3) all of the tests were different from one another and therefore the participants could not collaborate; and 4) the test administrator had no knowledge of which participants had which test kit [26].

Although both Bunch and Murphy's study and Giroux's study are referred to as "double-blind" (only Bunch and Murphy actually used this term; Giroux called his a "blind" study but acknowledges that he used the same experimental design as Bunch and Murphy), using the more generally accepted meanings of the terms "declared" and "blind" as described above, they could be more accurately characterized as "declared tests with two or more blind elements." While this may at first seem to be merely an exercise in semantics, the distinction is an important one to make when describing a test so as to be unambiguous regarding whether or not the participants had knowledge of the test. This description also fits the research project that is the subject of this article, since it follows a similar design.

One major difference between the validation studies cited above and proficiency tests like those provided by CTS is the fact that in the above studies none of the validation study test kits were exactly the same. This fact alone prevents the participants from being potentially biased by another nearby examiner also participating in the test. While the replicated test kits provided in CTS-like proficiencies and studies allow the researcher to assess variability in examiners' abilities, it does potentially allow for the inadvertent sharing of answers or outright collusion between examiners participating in the test. For example, if three examiners in a laboratory or laboratory system are taking a CTS test and Examiner 1 and Examiner 2 both identified Item 1 as having been fired in Gun 2, Examiner 3 can most likely conclude that Item 1 in their test kit will have been fired in Gun 2, if he or she somehow learns (either intentionally or unintentionally) of the other examiners' results.

Yet another factor that can be used in an experimental design to combat cognitive bias in declared firearm and tool mark identification validation and proficiency tests is the use of open test sets. An open set of test unknowns is one which may or may not contain a sample that matches one or more of the known samples provided to the participants. This is in contrast to a closed set, in which each of the unknown samples match one of the exemplars provided. If the participants know they are dealing with a closed set of unknowns, they will likely perform better on the test than if it were an open set because they may be able to use a process of elimination to infer at least a couple of the answers if they were able to identify the rest of the unknowns. Even if a test is designed using a closed set of unknowns, the test administrators can control bias to a significant degree by ensuring that no indication is given to the participants (via the test instructions) that they may have been given a closed set of unknowns [27]. This is what was done in the current study.

<u>The Scientific Method and Hypothesis Testing</u>

A common criticism of firearm and tool mark identification is that it is a subjective analysis that is not based in science or on the use of the scientific method [28, 29, 30, 31]. However, an analysis of the scientific method and its application in the field of forensic firearm and tool mark examination shows that the field is, in fact, well grounded in its use. Forensic firearm and tool mark examiners use the scientific method on a daily basis whether or not their educational background is in a physical science. Every case they work deals with each of the basic steps.

Any scientific inquiry usually starts with a basic question stemming from an observation. Questions asked in firearm identification typically include: "Did the submitted firearm fire the questioned bullet?" or "Were these two cartridge cases fired in the same (unknown) firearm?" The question then forms the basis of a hypothesis statement that is capable of being tested. This is an important point, because if a hypothesis cannot be tested (e.g., "Does God exist?"), then it is not something for which science can be used to find an answer. "Tested," in this sense, specifically refers to the *capability* of the hypothesis to be proven false through direct empirical observation (the null hypothesis). In science, certainty is an elusive goal. A hypothesis cannot be conclusively "proven" through empirical testing; it can only be disproven. If a hypothesis cannot be disproven, even over the course of years of repeated testing, the only implication is that its premise *may* be true, whereas once a hypothesis is disproven, it is rejected from further consideration. For this reason, a scientific hypothesis is usually stated in the negative form [32]. There is also another good reason for doing so, however. When forensic scientists deliberately state their hypotheses in the negative, they are taking precautions to minimize confirmation bias. In many instances, if a police agency submits a firearm and a fired bullet to the laboratory for an identification examination, the implied answer to the question they are asking of the examiner, "Was the submitted firearm used to fire the questioned bullet?," is that the submitter believes it did, otherwise the examination would not have been requested in the first place. The forensic firearm examiner can deliberately take a step back from potential sources of bias such as this by testing the negative hypothesis: "The firearm was not used to fire the questioned bullet." When approaching a case with this hypothesis in mind, the examiner proceeds with their examination in the normal way, but if this negative hypothesis cannot be disproven, then either the firearm was not used to fire the bullet or there is not enough information to disprove it, leading to an inconclusive result.

When a hypothesis is rejected, a new hypothesis is selected based on the results of the empirical testing. The process is then repeated to test the new hypothesis. The proper use of inductive and deductive logic is a critical component of this process. Grzybowski and Murdock have previously illustrated the application of inductive and deductive reasoning to the field of firearm and tool mark examination, as well as delineating the field's underlying scientific premises based on this reasoning [33]. When a hypothesis cannot be disproven and test after test continues to confirm it, then it can be considered a theory. However, as was true of the initial hypothesis testing, no amount of confirmatory test results can permanently or irrefutably establish the validity of a theory. For instance, throughout years of testing, the AFTE Theory of Identification has yet to be proven false. Yet, continued research and validation testing of the Theory, using the most current technology and philosophies in an effort to falsify the hypothesis that firearms and tool marks can be reliably identified on the basis of "sufficient agreement" of their microscopic characteristics, is needed in order to further develop and refine it. Limited testing of hypotheses, sometimes to the point of a researcher trying to prove a preconceived result, is one area that has been criticized by Mnookin, et al. in regards to forensic science; however, these authors also temper their criticism with some sound guidelines for forensic researchers:

> *Claims of knowledge should be taken as provisional and subject to revision in the face of new information. Dogma should be resisted. Research is not one thing, or one study, or once done, never reexamined. Research is an ongoing, incremental process. Research problems should be approached with an open mind. While it is certainly appropriate to have a hypothesis, or preliminary expectation, about what any given research study will show, investigators should follow the data whether or not it supports their original hypothesis, and whether or not it legitimates current practices. Research projects should be designed according to the norms of relevant academic fields. They should not be designed defensively, to produce, or to increase the chances of producing, a particular outcome* [34].

Case Influencing This Research

This research project was based on an actual officer-involved shooting case that was assigned to one of the author's colleagues. The firearms involved were two Smith & Wesson model 4006TSW, .40 S&W caliber semiautomatic pistols. As is common with these types of cases, one of the questions presented to the examiner was, "Which shots were fired from which firearm?"

In a firearm, a number of different parts are interacting with the cartridge case during cycling and firing. Each of these parts can be considered a separate tool, the working surface of which has to be evaluated individually. These different tools can leave either impressed or striated tool marks of varying levels of reproducibility, which is what makes cartridge case identification a unique challenge.

During the examination, the examiner noticed several details about the firearms that were of interest. First, the firing pins in these firearms are not fixed and can rotate during the firing process. If rotation occurs, it can cause the firing pin to present a different portion of itself to the surface of the primer during the creation of the firing pin drag mark. This different surface will create a different tool mark on the primer; therefore, the dragging of the firing pin may not create reproducible marks from one cartridge case to the next. The free-floating firing pin also made identification of the firing pin impressions more difficult because the impressions may not have been in the same orientation relative to the extractor and ejector marks on the fired cartridge cases from one shot to the next, creating the illusion of differences. In addition, the firing pins had very few surface defects on them and the resultant markings made by the two firing pins had some similarities.

Secondly, the manufacturing marks on the breech faces of the firearms had a linear parallel configuration, apparently from a broaching process, which indicated that subclass influences needed to be considered. This was compounded by the fact that the breech face marks did not impress well into the submitted evidence and ammunition (Remington Golden Saber™).

Thirdly, similarities in the ejector faces of both firearms were noted, indicating that these markings could again present a possible subclass influence.

The examiner who worked this case was ultimately able to use an ejector cut-out shear mark on the cartridge cases, along with the firing pin impressions, to make the final identification. The examiner indicated that this examination was more difficult than routine casework even though it was a closed set (or "known universe") in that the cartridge cases were known to have been fired in one of the two guns from the officers. The author felt this case example would be suitable for a declared blind study due to the fact that the surfaces were not marking well and the possible presence of subclass in the breech face marks.

In addition, the fact that the Smith & Wesson model 4006TSW is a relatively common firearm increases the practical value of the test. Not only do California Highway Patrol (CHP) officers carry the Smith & Wesson model 4006TSW, increasing the possibility that a California firearms examiner will have a case involving this type of gun, but it is also a model that has been available since 2000 [35].

This case prompted the question of how the presence of subclass characteristics may affect an examination. Two questions that arose from this case were: 1) In a declared blind study of cartridge case identification using samples (fired cartridge cases) in which subclass is possibly present, would the error rate of the participants be as low as currently estimated [36] or would the potential subclass influence cause an increase in the error rate?; and 2) Would the presence of potential subclass features in the test samples increase the chance of making an erroneous identification?

Taking these questions into consideration, the hypothesis for this study became: *When firearms examiners are presented with a no-gun examination, where subclass features may be present, and marks commonly used for identification transfer minimally to test-fired cartridge cases, they will not be able to determine that a particular cartridge case was fired in a particular firearm.*

Current Research

Three Smith & Wesson model 4006TSW firearms were obtained from the local CHP office. These firearms were from a pool of guns that were maintained as temporary replacement firearms for patrol officers in the event that their assigned firearm is sent for repairs or removed from service after an officer-involved shooting. Upon receipt, all of the working surfaces of the firearms that would come into contact with the cartridge cases, except for the chamber areas, were observed and documented using photography.

Breech faces:

On one of the firearms, the breech face had some very light, parallel, linear markings running vertically across the machined breech face. Based on their relatively uniform, uninterrupted appearance, some of these markings could potentially be subclass characteristics. The machined breech faces of the other two firearms also had vertical marks with an overall parallel linear configuration running from top to bottom; however, at the 12 o'clock position above the firing pin aperture the markings were intersected by other linear marks, forming an angular pattern. This angular pattern indicates that subclass should be less of an issue here because the positioning and relative spatial relationships of the intersecting markings appeared to be random. In addition, the linear marks on these breech faces also displayed an apparent random pattern of start and stop locations across the surface, which further minimizes the presence of subclass influence in these marks. **Figures 1 through 3** show silicone casts of the breech face markings of the firearms used in this study. The general reproducibility of the firearms used in this study can be seen in **Figures 10 through 12**, which show comparisons of the breech face casts to representative test-fired cartridge cases from each of the guns used. Based on this assessment of the breech faces, the presence of minimal subclass influences does not appear to be a significant problem in the potential identification of cartridge cases marked by them, at least within the limited population of these three firearms.

Firing pins:

The firing pins were removed from each of the firearms and inspected microscopically; see **Figures 4 through 6**. No

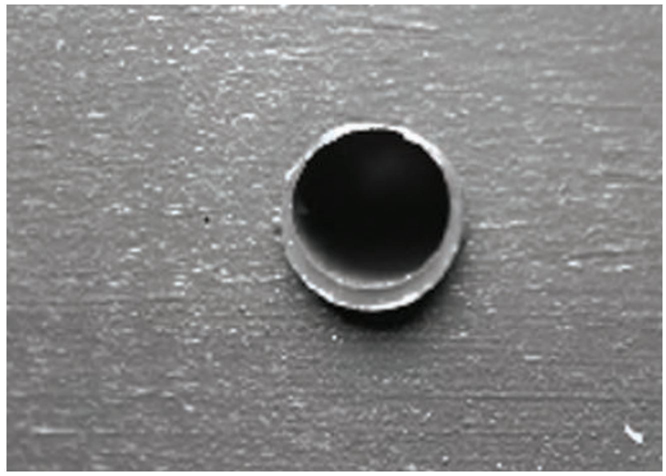


Figure 1: Gun 1, breech face cast, 20X [Marks appear horizontal in photo, but marks run vertically on original working surface of pistol]

Figure 2: Gun 2, breech face cast, 20X [same orientation as Fig. 1]

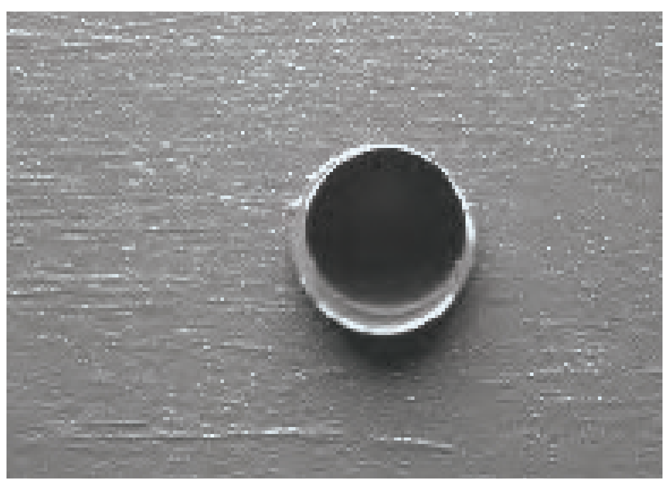


Figure 3: Gun 3, breech face cast, 20X [same orientation as Figs. 1 & 2]

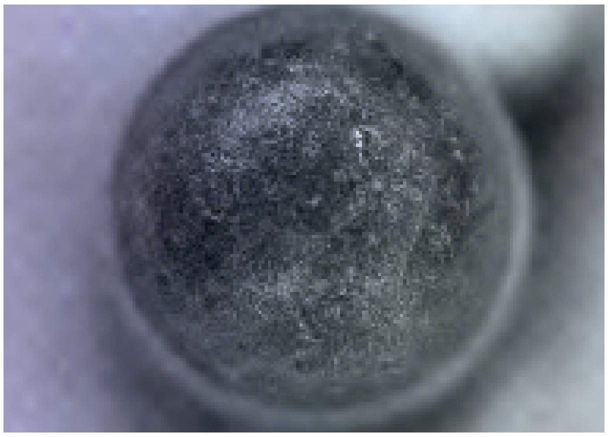


Figure 4: Tip of firing pin from Gun 1, showing multiple apparent random defects, 60X

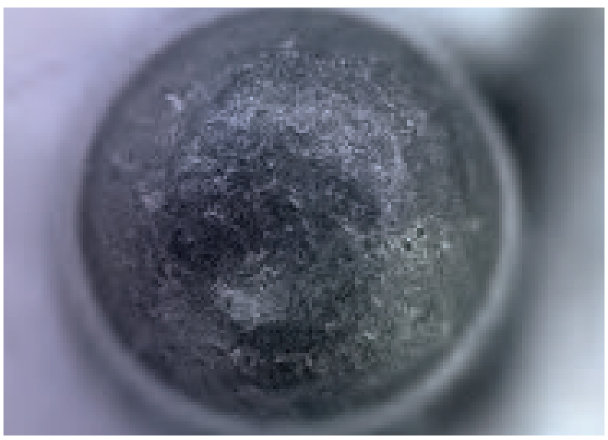


Figure 5: Tip of firing pin from Gun 2, showing defects, 60X

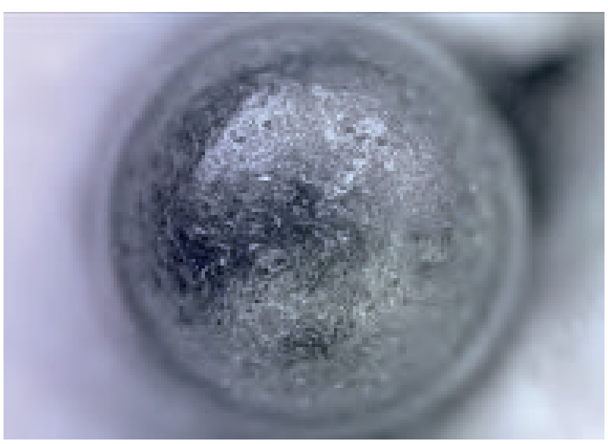


Figure 6: Tip of firing pin from Gun 3, showing defects, 60X

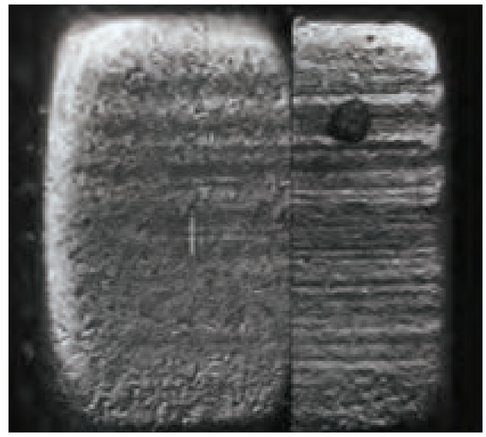


Figure 7: Apparent subclass carryover on ejector working surfaces; Gun 1 (left) vs. Gun 2 (right), 30X

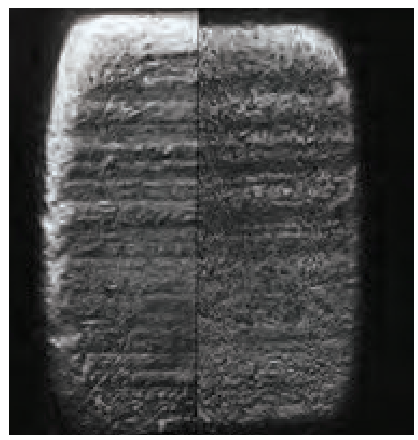


Figure 8: Working surfaces of ejectors, showing apparent subclass carryover; Gun 3 (left) vs. Gun 1 (right), 30X

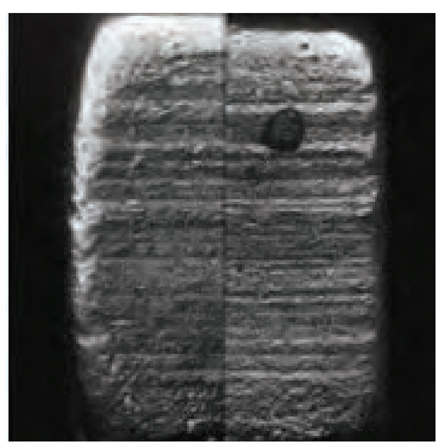


Figure 9: Working surfaces of ejectors, showing apparent subclass carryover; Gun 3 (left) vs. Gun 2 (right), 30X

circular lathe markings were observed on the firing pins to indicate that they were made using a lathe. Each of the firing pins had numerous surface defects in an apparent random pattern, which would indicate they were specific to these firing pins. These markings did not appear to be the same from firing pin to firing pin. However, because the firing pins have similar looking surface defects and are free to rotate during the firing process, an examiner could be confused by the random agreement found between impressions made by different firing pins.

Extractors:

The extractors for each firearm were examined microscopically. The surface of the extractor that would come into contact with the cartridge case could not be viewed directly due to its close proximity opposite the breech face, therefore the working surfaces of the extractors were not critically evaluated for subclass influence. In retrospect, casting would have provided an easy means of examining the working surfaces without having to remove the extractors for examination, or they could have been removed and examined. However, subclass characteristics have never been reported on the working surfaces (edges) of extractors.

Ejectors:

The faces of the ejectors were examined microscopically. Coarse striations across the ejector faces were observed. The ejectors from each firearm were compared to one another using a comparison microscope and some agreement of the coarse striae was found; see **Figures 7 through 9** for these comparisons. This indicated that there were subclass features present on these ejectors that may transfer to the cartridge cases during firing. This potential could present a problem to examiners if these features do in fact transfer to the cartridge cases and an examiner were to rely upon these marks for identification without evaluating them for subclass.

Chamber marks:

During the initial evaluation of the firearms, the chamber marks were not examined and the chambers were not evaluated for subclass influence. After the results from the test kits were received it was found that a number of examiners used the chamber marks to aid in their identification. Since the working surfaces of the chambers had not been examined prior to distributing the test and by this time the original three firearms had already been returned to the CHP and were not available for reexamination, yet a fourth Smith & Wesson model 4006TSW firearm was obtained from a retired CHP officer in order to examine these surfaces. A silicone cast of

Figure 10: Gun 1, breech face cast (left) vs. breech face marks on fired cartridge case (right), 30X

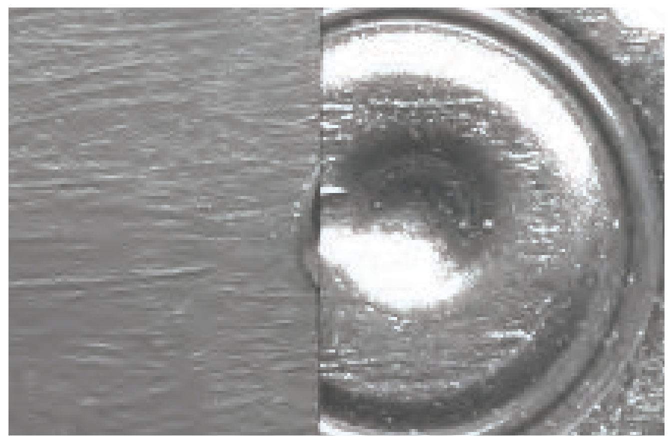


Figure 11: Gun 2, breech face cast (left) vs. breech face marks on fired cartridge case (right), 15X

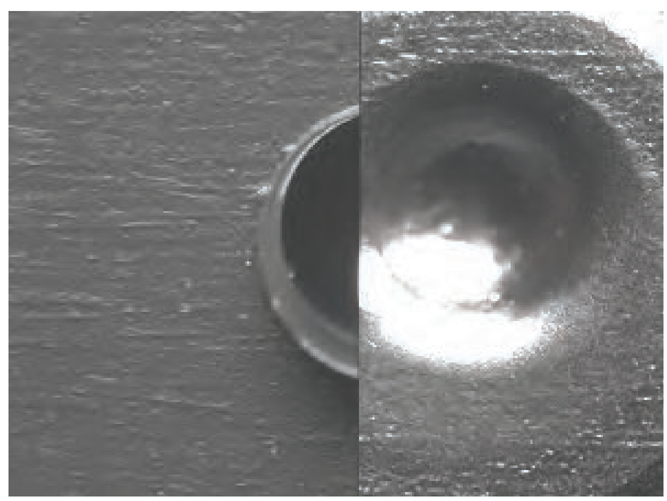


Figure 12: Gun 3, breech face cast (left) vs. breech face marks on fired cartridge case (right), 30X

the chamber area was made and the cast was examined using a stereomicroscope at various magnifications. The chamber marks appeared to be very small chatter-type marks distributed longitudinally around the entire circumference of the chamber surface. In this particular firearm, there was an area of the chamber that appeared to have a raised metal defect that would most likely transfer an impressed and/or striated mark to a cartridge case. The chatter-type marks appeared to be random and were oriented parallel to the direction that the cartridge case would enter and exit the chamber. Despite being parallel to the direction of the cartridge case insertion and extraction, their apparent random nature should make them unique characteristics for purposes of identification. Therefore, if the three original pistols had a similarly random array of chamber tool marks as the fourth firearm appeared to have, they too should have had no subclass influence. The author feels this is a reasonable assumption considering these firearms were batch produced using the same tools and machining processes.

Information on How These Working Surfaces Are Manufactured

A literature search was conducted for information on how the working surfaces of the Smith & Wesson firearms that would potentially be used for this study were tooled. In 2010, Lightstone provided a good summary of how some of these surfaces were made [37]. According to the information Lightstone received in regards to Smith & Wesson Sigma series firearms at the time of her Smith & Wesson factory tour in June 2009, the breech faces are cut in a single pass of a horizontal broach. After broaching, the firing pin aperture, extractor hole, and the firing pin assembly hole are drilled into the breech face using a gun drill. This drilling process leaves burrs on the surfaces which are later hand-sanded off. The slides are tumbled in a ceramic medium that has potential to contact the breech face if the media is small enough. The slides are then sand-blasted to smooth the outside surfaces, but the abrasive material often contacts the breech face surfaces as well. The final finishing step is glass bead blasting the slides. This bead blasting often contacts the breech faces of the slides even though it is not directly aimed at them.

Lightstone's article also indicated that Smith & Wesson was going through a transition from the above described process to a newer, more automated process using fewer computer numeric control (CNC) machines. The main difference in this new process is that the breech faces are cut using two separate broaches. The first broach flattens out the rough areas and is followed by a step-broach which finishes the surface [38].

According to weapons officer Scott Fredrick of the CHP's Northern Division office, they received their first lot of Smith and Wesson model 4006TSW firearms in December 2006 and their order was completed in October 2007. The fact that Smith & Wesson did not switch over to its new CNC machining method until 2009 would indicate that the CHP firearms used for this study were produced using the older method of manufacturing.

In May 2011, the author was given the opportunity to tour the Smith & Wesson factory and ask questions of Product Engineer Jason Dubois regarding the machining of some of the parts used in model 4006TSW firearms. He indicated that the model 4006TSW, which is similar to the civilian 3rd Generation models, has been phased out of production but is still available to be purchased in a batch quantity if an agency were to request it. While some of the manufacturing techniques from the time the model 4006TSW's were produced have changed, the basic processes for their current firearms are as follows:

For the newer models, the M&P series and the SD series, the slides start as a piece of bar stock of the desired metal. The bar stock is placed into a CNC machine which will perform five separate operations on the bar stock to bring it to its final shape and dimensions. The breech face of the slide is broached inside the CNC machine with a rectangular-shaped step broach. Once the slides are taken off of the CNC machine, they are wet blasted with a gritty material to de-burr any of the rough surfaces created during cutting. The slides are then sent for heat treating. If the slide is to be coated with Melonite®, they are sent to an outside vendor who performs this operation and then returns them to Smith & Wesson. The switch to a single CNC machine, versus the old production method which required more than one machine for all of the cuts, requires the use of fewer employees and also allows for less error in moving the parts from machine to machine. The Sigma series slides were produced in the older method described in Lightstone's paper, which utilized a manual horizontal broach. The Sigma series pistols have since been discontinued.

The tools used on the CNC machines have a designated "tool life" that is programmed into the machines and is based on minutes of cutting time. Once a tool has reached its maximum amount of cutting time, the machine automatically switches out the tool and flags it as being in need of sharpening. If the tool is designed for precision cutting, the tool life is much shorter between sharpenings. For rough cutting tools, the tool life can be extended for a greater amount of time. Whether or not the tool will be re-sharpened or just replaced depends on the cost, design, and age of the tool. All of the tools have an acceptable limit to which they can be resurfaced, and once they have reached a point beyond this tolerance they are replaced.

During production of the slides, a certain percentage of them will be removed from the production line and their dimensions measured using a coordinate measuring machine (CMM). This is done to determine if the slides are being cut to the appropriate dimensions and shape by a particular CNC machine. This process is done on the manufacturing floor in order to allow any mistakes to be caught early in the production process.

For all of the pistols made by Smith & Wesson, the chamber surface is honed using a ball brush. A ball brush is a wire brush that has small balls of grit attached to the bristles. The grit material can either be metal or nylon depending on the surface that is being honed. The ball brush is attached to a machine that rotates the brush at high speeds, similar to a drill. The brush is then inserted into the chamber area and the grit smooths out the inside of the chambers.

The firing pins used in Smith & Wesson's firearms are manufactured in a number of ways, depending on whether they are made in-house or are purchased. If they are made in-house at Smith & Wesson, they are turned on a screw machine and then placed into a ceramic medium where they are tumbled. If they are purchased from an outside vendor they are generally a metal injection molded (MIMed) part that is then tumbled or coated.

The extractors used by Smith & Wesson are all purchased from outside vendors, except for the 1911 series. The purchased extractors are MIMed and have no finishing performed by Smith & Wesson; any finishing processes on these extractors would have been performed by the outside supplier before Smith and Wesson received them.

As with the firing pins, the ejectors are manufactured in a number of ways: MIMed, stamped, or fine blanked. The MIMed ejectors are only used in the 1911 series and are purchased by Smith & Wesson. The stamped ejectors are also purchased from an outside vendor. The fine blanking is done by Smith & Wesson in their Maine factory. These fine blanked ejectors are the ones used in the model 4006TSW pistols. The ejectors used in Sigma series pistols were fine blanked by a vendor. Six of the fine blanked ejectors were provided as production samples by Smith & Wesson in order to evaluate their surfaces. On close examination of these samples, it was evident that there were subclass characteristics present on the ejector face. The presence of similar subclass on the ejectors of the Smith & Wesson model 5906 was noted by Evan Thompson in 2002 [39]. Thompson contacted Smith & Wesson and was informed that these ejectors were stamped from bar stock and then the faces were ground. Not only did the casts Thompson made of the ejectors display very similar markings as the marks seen on the ejector samples provided to the author by Smith & Wesson, but they appeared very similar to the ejectors seen in the firearms from the real case example and the three firearms used in this study as well. If the ejector faces were ground, as Thompson was led to believe, subclass carry-over would be very unlikely based on the rapidly changing surface of most grinding wheels. However, firearm examiners are aware that depending on the manufacture and material of the grinding wheel, the hardness of the material being ground, the wear rate of the grinding wheel, and the fact that the face of the ejectors are very small surfaces, it is possible that subclass characteristics could carry over from one ejector to the next even if a grinding wheel was used [40]. The fact that these ejectors are reportedly fine blanked makes subclass more of a possibility due to the fact that the same die will be used to make each of the parts and the dies will change much less rapidly over time than a grinding wheel. Dubois did not know if any further finishing was performed on these parts but indicated that due to the precision machining techniques now used to make them, very little hand fitting is required on the final products.

Given the manufacturing techniques described in Lightstone's paper and those observed by the author during her own Smith & Wesson factory tour in 2011, the potential for subclass characteristics is present on the breech faces and ejectors of model 4006TSW firearms and any other series of Smith & Wesson firearms that are produced using the same manufacturing techniques.

Materials and Methods

Independence brand, .40 S&W, 180 grain, FMJ ammunition was purchased in as few different lots as possible in order to limit the amount of variables in the research study (according to several internet sources, this brand of ammunition is made at the ATK-operated Lake City Army Ammunition Plant in Independence, Missouri, using Federal, CCI, and Speer components). A local gun store was able to provide enough ammunition for the study in two lots: A22R39 and A23R31. Prior to creating the test kits, the head of one cartridge from each box of each lot was examined and photographed using a Leica FSC comparison microscope at various magnifications to determine whether or not any manufacturing marks could be potentially mistaken for firearm-produced marks. While numerous small, irregular manufacturing marks (produced either during or incidental to the manufacturing process) were apparent on the case heads, none of these appeared to be in a quantity or configuration sufficient to interfere with comparison. Just to be absolutely certain, a Smith & Wesson model 4006TSW firearm was obtained and 10 of the Independence brand cartridges were fired through it and the cases were collected. These cases were examined using a comparison microscope and the manufacturing marks on the cartridge cases were found to either be obliterated during the firing process due to obturation and impressing of the breech face and firing pin, or were easily distinguished from the breech face marks on the primer. Therefore, these pre-existing marks would not likely be confused by an examiner as having been produced by the firearm during firing.

The three Smith & Wesson model 4006TSW firearms obtained from the CHP for this study were designated “Gun 1,” “Gun 2,” and “Gun 3,” respectively, and the serial number for each firearm and its designation were documented. The serial numbers for the three firearms were within 300 numbers of one another and two were within 20 numbers of each other. While these numbers being close in sequence does not necessarily mean that the firearms were actually produced sequentially in this same order, it does indicate that they were made within a relatively short time frame of one another.

Three magazines were provided by the CHP with the firearms. In order to speed up the test firing process and to reduce the amount of time necessary to stop and reload magazines, all three magazines were used during the test firing of each firearm. Because the same magazines were used throughout the test firing, the magazine lip marks could not be used by an examiner during the examination process. The examiners who participated in the study were instructed to not use the magazine lip marks in their comparisons.

<u>Known Cartridge Cases:</u>

Each firearm was used to fire 124 cartridges. Ammunition lot number A23R31 was used for Gun 1, while lot number A22R39 was used for the test-fired components fired in both Gun 2 and Gun 3. Each firearm was used to fire 93 cartridge cases to be used as known specimens. These cartridge cases were collected and engraved with the gun number inside the mouth. The remaining 31 cartridge cases from each gun that were set aside as unknown, or questioned, were not engraved at this time but were placed into a box labeled with the firearm number in which they were fired. These unknowns were labeled with a random identifier by another examiner at a later time in order to keep the study blind in this respect.

Three known cartridge cases from each firearm were placed into each of 31 test kits that were created. **Figures 13 through 18** show representative comparisons of the breech face marks and firing pin impressions from each of the three firearms on cartridge cases provided in the test kits.

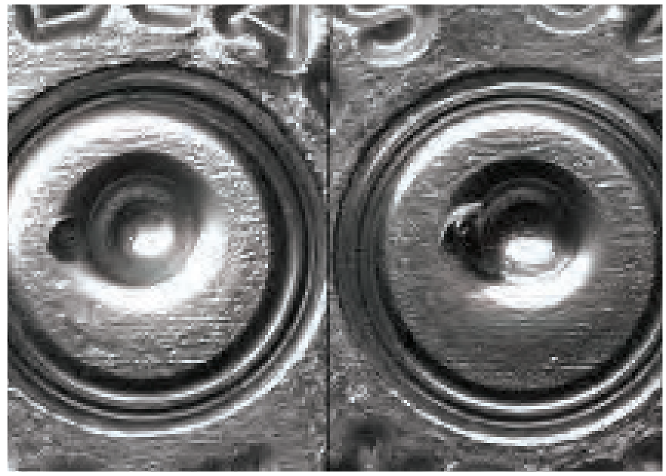


Figure 13: Gun 1, reproducibility of impressed striae in breech face impressions on two representative test-fired cartridge cases, 10X

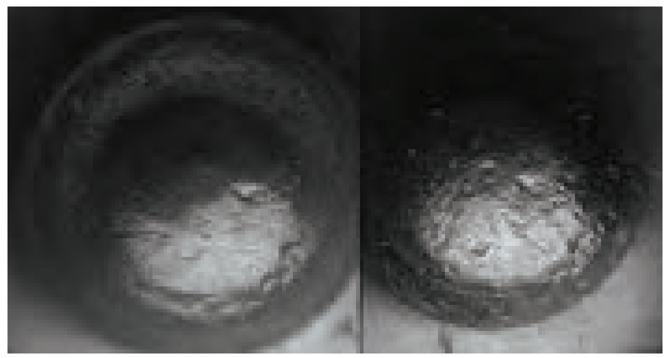


Figure 14: Gun 1, reproducibility of impressed marks in firing pin impressions on two representative test-fired cartridge cases, 30X

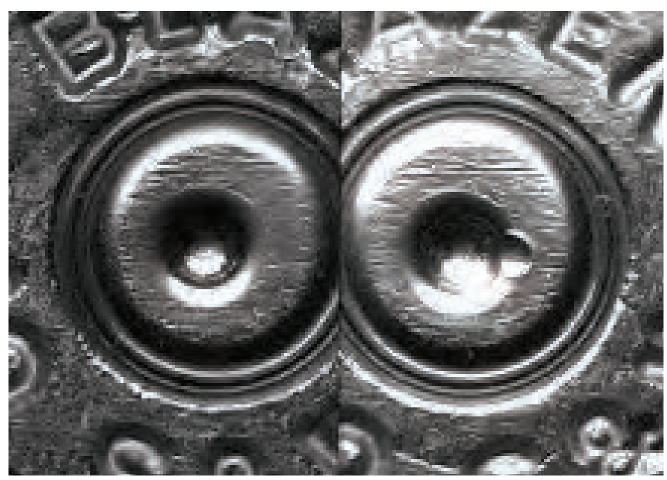


Figure 15: Gun 2, reproducibility of impressed striae in breech face impressions on two representative test-fired cartridge cases, 10X

Unknown Cartridge Cases:

The three boxes labeled "Gun 1," "Gun 2," and "Gun 3," containing the unknown cartridge cases, were given to another examiner in the laboratory who then engraved each cartridge case with a randomly generated number. The numbers used were generated from a random number generator program available on the internet. The second examiner documented which randomly numbered cartridge case was associated with each firearm prior to mixing all of the unknown cartridge cases from all three firearms in a cardboard box.

A third examiner randomly selected three cartridge cases from the box for each test kit. The numbers selected for each kit were recorded, along with the kit number that they were assigned to. With each stage of the kit-making process performed by different examiners, no one examiner was aware of how a particular kit was produced.

The examiner who performed the original test firings then randomly selected the assembled kits for mailing out to the external examiners (one kit per examiner) who agreed to participate in the research project. In total, 30 kits were sent out to a total of 14 different crime laboratories. Multiple kits were sent to a few of the participating laboratories in order for more than one examiner in those laboratories to participate in the study. Agencies from eight different states in the U.S. and one from Canada participated in this study.

A standardized answer sheet was provided with each kit. The answer sheet asked for very basic demographic information from the participating examiners, including: number of years in the field, approximate case load per year, if the examiner's laboratory was accredited, and if the examiner was certified. The participating examiners were not only from very different geographic locations but also had very different experience levels. The number of years in the field ranged from 2.5 to 30 years with the average number of years in the field being 13. The number of examinations worked per year for the participants ranged from 20 to 500 with the average number being 150 examinations. There were two outliers in this demographic information, including one participant who reported that they no longer perform firearm casework and another who reported they performed over 1000 examinations per year if each bullet or cartridge case examined was considered to be a single examination. All of the participants in the study worked for accredited laboratories. Eight of the participants were certified through either AFTE, the American Board of Criminalistics (ABC), or both. One of the examiners stated they were "certified" by "FirearmsID" (although it is unknown exactly what this examiner was referring to with this response, the website www.FirearmsID.com offers written

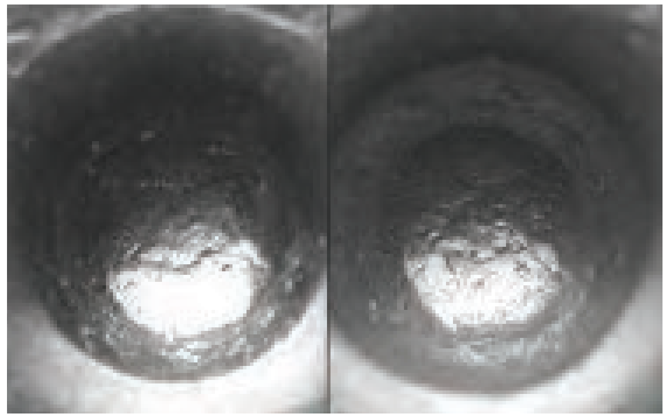


Figure 16: Gun 2, reproducibility of impressed marks in firing pin impressions on two representative test-fired cartridge cases, 30X

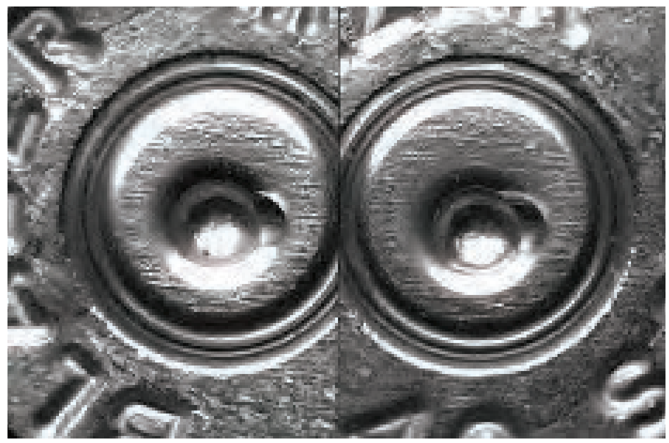


Figure 17: Gun 3, reproducibility of impressed striae in breech face impressions on two representative test-fired cartridge cases, 10X

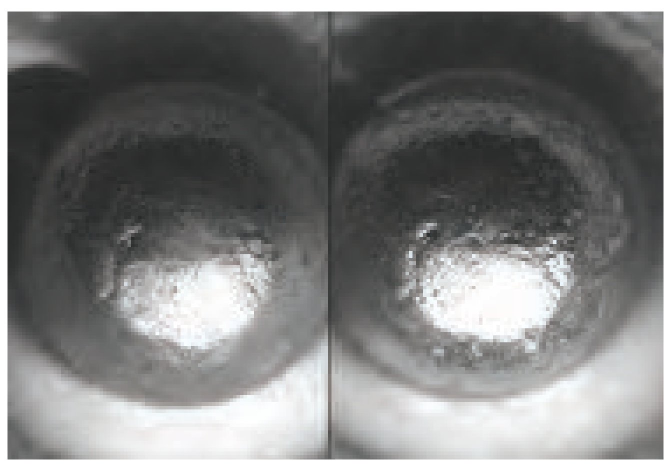


Figure 18: Gun 3, reproducibility of impressed marks in firing pin impressions on two representative test-fired cartridge cases, 30X

and practical self-tests in firearms identification to interested users, but this is not an independent certification program).

Throughout the entire preparation of the test kits, no individual preparer had access to the answer key for any of the test kits. For example, the person who made up the known cartridge cases did not know what the unknown cartridge case numbers would be for each firearm. The person who designated the unknown cartridge case numbers did not know which test kit these unknowns would be associated with or from which of the known firearms they came. The final kit preparer did not know which firearm any of the unknown samples came from nor did they know the serial number of the firearm for the knowns or the unknowns. Lastly, the person who sent out the test kits only knew from which serial number firearm the knowns came and was not privy to any of the answers for the unknowns.

Similar to the Bunch and Murphy and Giroux studies, an element of anonymity was provided to the participants in this study. A list of the kits and the agencies to which they were mailed was maintained, but only for tracking purposes and it was not retained after the kits were returned. This list was kept in the possession of a member of the clerical staff at the laboratory and was not revealed to the test creators. This was not a study about how an individual examiner or any particular agency would perform; it was intended to survey a cross-section of the firearms examiner community. Because the test participants were well aware they were participating in a test, and the aforementioned precautions were taken in order to blind the test administrators to any bias their knowledge of the test results might produce, this study can be considered to be a declared test with several blind elements.

The examiners were asked to examine the items as they would work a normal case, including second calls or technical reviews, if authorized by their laboratory procedures. The questions the examiners were asked to answer were: 1) “Were any of the unknown expended cartridge cases discharged by the same firearm as the known expended cartridge cases from Gun 1?”; 2) “Were any of the unknown expended cartridge cases discharged by the same firearm as the known expended cartridge cases from Gun 2?”; and 3) “Were any of the unknown expended cartridge cases discharged by the same firearm as the known expended cartridge cases from Gun 3?”. Although the test kits were designed so that each of the unknown cartridge cases had a corresponding known specimen in the same kit (a closed set), the questions were worded in a manner that would not imply to the participants that they were dealing with a closed set. The examiners were asked to answer these questions using one of three conclusions as defined by the Range of Conclusions in the AFTE Glossary: 1) Identification, 2) Inconclusive[1], or 3) Elimination [41]. In addition to these questions, the participants were asked what areas of the cartridge cases they relied upon for their conclusions. Without directing the examiners as to what marks to use, the answer sheet provided the following choices: “breech face”, “firing pin impression”, “extractor”, “ejector”, “chambering marks”, and “other”. As previously mentioned, the participants were instructed not to use magazine lip marks since the same three magazines were used to fire all of the knowns and unknowns.

Results

Of the 30 test kits that were sent out to the participating agencies, results from 25 of the kits were returned. Since three unknowns were provided in each kit, this equals a total of 75 results. Of these, 74 of the responses correctly identified the firearm that had fired the unknown cartridge cases. One result was reported as inconclusive. This examiner did not indicate what led them to their inconclusive result (See addendum at the end of this paper).

Of the 74 correct responses, breech face marks (BFM) were indicated to have been used, at least in part, to reach the reported conclusions in all of them. Twenty of these responses stated BFMs were used exclusively, while an additional three responses implied that only BFMs were used through comments such as, “Used BFM for all identifications however, did not use gross contours of BFM (only used those to phase) – used smaller/individual w/in cross contours to ID,” and “Gun 2 had excellent fine detail in breech face impressions.” One of these examiners indicated that they were concerned that the parallel breech face marks may have been subclass so they looked for “endings, chatter, finer marks within, and nicks that appeared to be from other sources.” These comments indicate that although parallel breech face marks have the potential for subclass features to be present, at least some firearm examiners (the ones who commented) are aware of this and know how to evaluate the surfaces properly to limit the effect that subclass might have on their examinations. It should be noted that the examiner who reported the inconclusive result was the sole participant who did not indicate they used the breech face marks.

The second most relied-upon area for the participating examiners was the firing pin impression. The examiners indicated in 46 of the examinations that they used the firing pin impressions in part for their identifications. One examiner

[1] For the sake of simplicity, the three subcategories of inconclusive results and the conclusion of “Unsuitable”, as described in the AFTE Range of Conclusions, were not offered as possible response choices; however, it is acknowledged that in actual casework the nature of any inconclusive result should be further described.

indicated that the firing pin impressions in the test fire sets "did not demonstrate sufficient repeatability to use that tool mark for identity or elimination." No other comments were made in regards to the lack of reproducibility of the firing pin impressions, but at least two examiners indicated that they cast the case heads to better visualize the firing pin impressions. It is possible that the nickel-plated primers of these cartridge cases prevented good visualization of the firing pin impressions. The fact that the firing pins in the firearms used to prepare the tests had a "floating," or non-fixed, design could have made the resulting impressions difficult to use for identification due to their lack of consistent orientation in reference to the firing pin drag marks or extractor and ejector locations.

In five responses, the participating examiners indicated they also used the ejector mark in their examinations, in combination with either the breech face marks only (three responses) or with the breech face marks and extractor marks (two responses). Some of the examiners indicated that they would not rely on the extractor or ejector marks for identification without having the gun to examine because they could not ascertain how these parts were made or evaluate them for the possibility of subclass features. However, a reading of the available literature on this subject shows there is a higher potential for subclass influence to be an issue with ejectors than it is with extractors.

A study by Nichols described the possibility of subclass characteristics on the non-working surfaces of extractors [42]. While the particular surfaces Nichols described as exhibiting subclass in his study did not make contact with the cartridge cases during firing or ejection in his study, the comments from some of the participants in this study show that firearm examiners are wary of the possibility of subclass characteristics on extractors, even though subclass has never been reported on their working surfaces (edges). Conversely, subclass is known to be a potential problem on the forward faces of ejectors, as indicated by Thompson's above-mentioned article regarding the Smith & Wesson ejectors he encountered. However, as has been previously reported, the possibility of such characteristics getting transferred to the cartridge case depends greatly on the orientation of the ejector face to the head of the cartridge during cycling of the firearm [43]. The ejectors of the firearms used in this study did, in fact, display potentially significant subclass carryover on their working surfaces at 30X magnification (**Figures 7 through 9**), but it is not fully known to what extent, if any, this subclass influence carried over to the marks they left on fired cartridge cases. Based on the microscopic intercomparison of a few of the ejector marks at 120X magnification on fired cartridge cases produced by the firearms used in this study, reproduction of the subclass influences present on the ejector faces does not appear to have been a significant factor, because the detail seen at 120X is smaller than the detail observed at 30X. Even if some subclass characteristics were imparted to the ejector marks, those examiners who stated that they used these marks as a basis for identification did not indicate to what extent they relied on them to reach their conclusions and could have merely used them as a class characteristic for orientation purposes due to their relative position on the cartridge case. Generally speaking, extractor and ejector marks on fired cartridge cases may not be as routinely compared as breech face marks and firing pin impressions are in firearms identification casework involving semiautomatic pistols, since the identification of extractor and ejector marks typically cannot support a conclusion that a cartridge case was fired in a gun like true cycle-of-fire marks (e.g. breech face marks and firing pin impressions) can. Still, reporting a finding that a cartridge or cartridge case was cycled through the action of a particular gun on the basis of extractor and ejector markings can be useful information in an investigation.

In eight of the reported examinations in this study, the chamber marks were used to aid in the identification. In seven of the examinations, it was reported that "other" features were also used. Some of these other features were indicated to include: "ejector cutout," "unknown-possible chambering mark," and "firing pin drag."

Discussion

This project was designed to provide information on error rates in firearm related casework in a study that presented as little bias to the participating examiners as possible. In order to achieve this goal:

- The participants in the test were unaware of the answers.
- The kits were made up in a random fashion in an effort to prevent any two tests from being exactly the same. The test was designed so multiple examiners in one laboratory could participate in the study; therefore, each kit had to be different to prevent participants from possibly influencing each other's answers.
- The test administrator was not aware of the answers for any of the test kits, as these were furnished by another examiner in the laboratory.
- The kits given to each examiner were documented only for tracking purposes and the list was not retained once the testing period was completed. At no time was this information viewed in relationship to answers received,

but only in an effort to locate missing test kits.

- The examiners tested were asked to treat these test kits as they would treat casework; second calls ("verifications") and technical reviews were allowed and encouraged.

While the answers to each test kit were documented and retained in order to determine if the participants came to the correct conclusions, it was understood that errors in kit construction could occur. An effort was made to ensure that the unknowns in each kit were properly documented and that there were no transcription errors. Despite this precaution, the possibility still remained that an error could have occurred. The author was prepared to investigate any submitted result that was incorrect to ensure that this was an error made by the examiner being tested and not due to an improperly designed or documented test kit.

It is well understood among experienced firearms examiners that the same firearm can mark fired cartridge cases differently from shot to shot, even using the same ammunition, due to any number of variables that may occur during the firing process. For example, a firing pin may impress deeper into the primer in some tests than others, breech face marks may be visible on some cases while not present on others, an ejector may mark in the lettering of the headstamp and therefore not be discernible on some cases while marking others in a much clearer manner, and firing pin drag may occur only a fraction of the time. It is because of these variables that inconclusive results were not unexpected.

The firearms used in this study were known to produce similar markings to one another, which potentially made the examinations more challenging. This study was designed to present real world casework to examiners in a format that could be used to determine a realistic frequency of error in a real world situation. While the types of firearms used for this study may be more commonly encountered by firearm examiners in California than in other parts of the United States, this scenario mimics any officer-involved shooting scenario where the majority of the cartridge cases located at the scene will be from firearms of the same make and model. For many firearm examiners, their local law enforcement officers will have been issued Glock firearms [44]. A shooting case involving Glock-fired cartridge cases generally represents a best-case scenario for any firearm examiner because of the relative ease with which the firing pin aperture shear marks on these cartridge cases can be identified. However, Smith & Wesson firearms, including M&P or other third generation type pistols, are generally the second most issued firearms to law enforcement. From information provided by Smith & Wesson during the author's factory tour, the slides for these firearms are produced similarly, so it can be inferred that they will all likely produce tool marks on cartridge cases in a fashion similar to those of the Smith & Wesson model 4006TSW firearms used in this study. However, this assumption was not tested by the author and may be an interesting area for additional research.

Despite the difficulties that were specifically designed into this study to present the participants with realistic samples inspired by an actual case, all of the participating examiners were able to correctly distinguish which cartridge case was fired in which firearm, with only one examiner reporting an inconclusive result. Given these results, the author must *reject* her original hypothesis. This hypothesis stated that when firearm examiners are presented with a no-gun examination, where subclass features may be present, and marks commonly used for identification transfer minimally to test-fired cartridge cases, they will not be able to determine whether or not a particular cartridge case was fired in a particular firearm.

Areas for Additional Research

Looking ahead to future research studies that may deal with similar subjects, a more thorough pre-screening of any samples designed to test examiners' assessment of subclass features is obviously one area of the current study that could be improved. This study could also be further improved upon if additional firearms were used and knowns from only a portion of those firearms used in the test kits, thus presenting an open set of unknowns to the participants. While this could increase the chances of inconclusive results, it would be a more accurate reflection of the types of evidence received in real casework. With regard to the reporting of inconclusive results, it would have been better to direct the participants to specify which category, 2. a), b), or c), as described in the AFTE Range of Conclusions [45], best described their observations, rather than merely reporting "inconclusive". Had this been done, it may have provided some insight into what the participant who reported the inconclusive result observed that would not allow them to either identify or eliminate that particular unknown. Requesting that participants submit photographs of any comparisons that result in identification or inconclusive opinions could also go a long way towards answering any lingering questions (especially in instances where potential subclass characteristics may be an issue), but this would also require more time on the part of the participant.

Another way in which this study could be improved would be to use firearms having more prominent subclass features that would transfer to the fired components more readily. While some subclass characteristics appeared to be present on some of the working surfaces of the firearms in this study, most did not transfer significantly to the cartridge cases presented to the participants. The use of sequentially-manufactured firearms having more defined subclass features would present a greater challenge to the examiners participating in the test.

Addendum: The Single Inconclusive Result

The author is of the opinion that the sole inconclusive result reported for this test was likely the result of cognitive bias on the part of the participating examiner. In order to explore the issue, the test kit that had been assigned to this examiner (who reportedly had 13 years of experience in firearms identification) was re-examined by the author in order to assess why the examiner may have come to this conclusion. Upon viewing the cartridge case in question, Unknown 2, on the comparison microscope, it was apparent that the case had good, clear markings on it. The cartridge case lacked a firing pin drag mark but it had clearly discernible breech face marks and a fairly deep, well-marked firing pin impression. Because the correct results were known to the author, the cartridge case was compared to one of the known cartridge cases from Gun 1, which had been documented on the answer key as having fired Unknown 2 for this kit. The observed agreement in the firing pin impression and the breech face marks was sufficient in quality and quantity to identify Unknown 2 as having been fired in Gun 1. The answer sheet for this examination was re-evaluated and it was noticed that the examiner had marked the result as "inconclusive" when comparing Unknown 2 to Gun 3. There was no indication on the answer sheet as to whether or not the examiner had eliminated, identified, or found inconclusive findings for this cartridge case in comparison to Gun 1 and/or Gun 2. It is interesting to note that in this particular test kit, Unknown 1 had been fired in Gun 2 and Unknown 3 had been fired in Gun 1, both of which the examiner correctly identified. Unknown 2 (the reported inconclusive) had also been fired in Gun 1. It is possible that the examiner assumed that corresponding unknowns were provided for each of the known firearms represented in the test kit, since each kit contained three unknowns and three knowns. If this assumption was made, the examiner could have thought Unknown 2 was likely fired from Gun 3 (since it had not been identified to either of the other unknowns) but could not find sufficient agreement between the compared markings to support this conclusion, hence the inconclusive result. However, this is speculation on the author's part and this inconclusive result could be due to any number of other factors, including fatigue or time constraints.

In order to assess what the participating examiner may have observed during their examination of this unknown that led them to report an inconclusive result, the author compared the unknown cartridge case to the known cartridge cases from Gun 3. Random agreement, of the type that can usually be found in the comparison of known non-matching tool marks, was found in both the breech face marks and in the firing pin impressions, but not in sufficient quantity to lead a trained examiner to misidentify it. A comparison of Unknown 2 with Unknown 3 (both of which, for this kit, had been fired in Gun 1) showed there was sufficient agreement of the firing pin impressions for identification; however, there was insufficient agreement of the breech face marks. While this information was useful, the author was aware of the fact that she was potentially subject to confirmation bias because she knew the true answer. Therefore, a second qualified firearm examiner in the author's laboratory was presented with these items for examination without being given any contextual information. The second examiner compared a single known cartridge case from Gun 3 to Unknown 2 with inconclusive results. The examiner subsequently identified a known cartridge case from Gun 1 to Unknown 2. See **Figures 19 through 24**, which were taken during the re-evaluation of this test kit.

Acknowledgements

The author would like to acknowledge the following people for their contributions to this research: Michael Barnes of the California DOJ Redding Regional Crime Laboratory; the NFEA class of 2011 and staff, including S.A. Jim Yurgelitis, Jodi Marsanopoli, Sheila Hopkins, and all of the instructors; all of the examiners who participated in this study or offered to have their employees participate in the study; Jason Dubois of Smith & Wesson; and the California Highway Patrol.

The author also gratefully acknowledges and appreciates the significant editorial contributions made by Eric Collins and John Murdock to the blind trial and scientific method portions of this article. And finally, but of the utmost importance, although the research and methodology employed were planned and performed by the author, this paper, with the consent of the author, has been heavily edited and revised by Eric Collins. The author is in his debt for doing so.

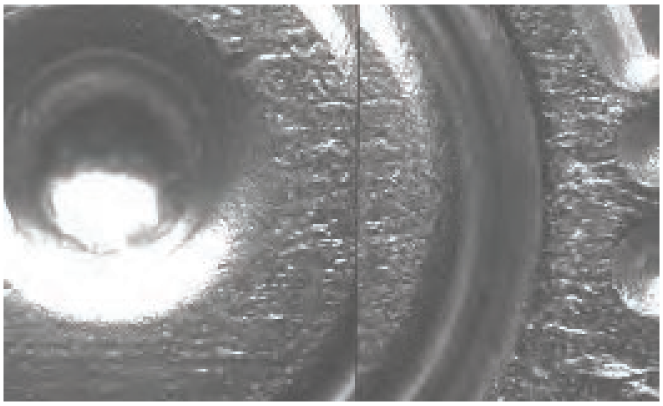


Figure 19: Comparison of breech face marks on test-fired cartridge case from Gun 1 (left) vs. Unknown Cartridge Case #2 (right), showing sufficient agreement for identification, 30X

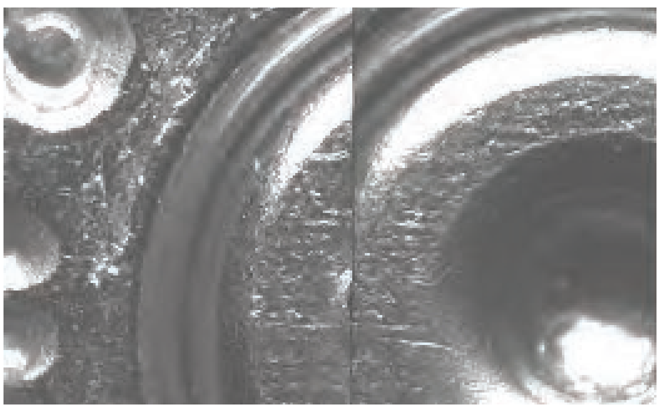


Figure 20: Comparison of different area of breech face marks on same items as Figure 22, showing sufficient agreement for identification, 30X

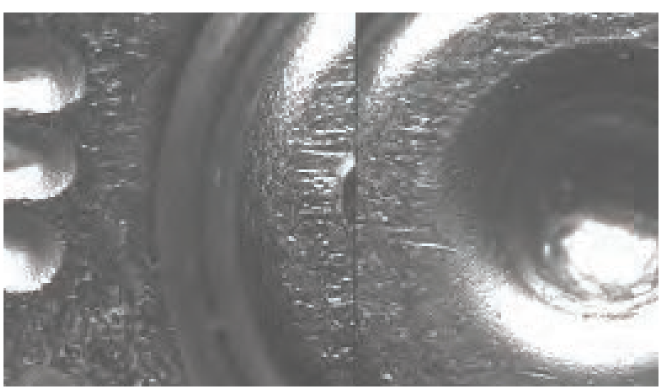


Figure 21: Comparison of breech face marks on test-fired cartridge case from Gun 3 (left) vs. Unknown Cartridge Case #2 (right), showing insufficient agreement for identification, 30X

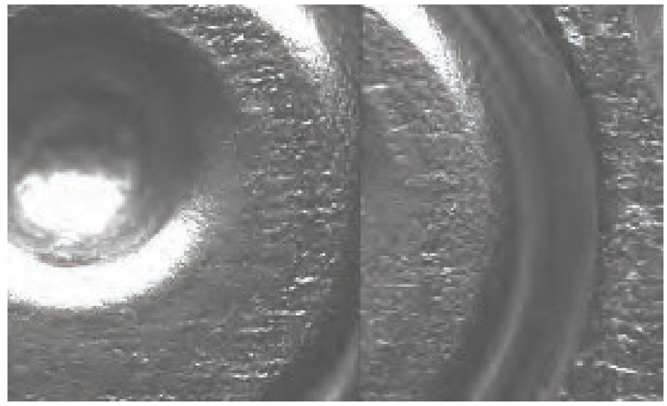


Figure 22: Comparison of different area of breech face marks on same items as Figure 22, showing insufficient agreement for identification, 30X

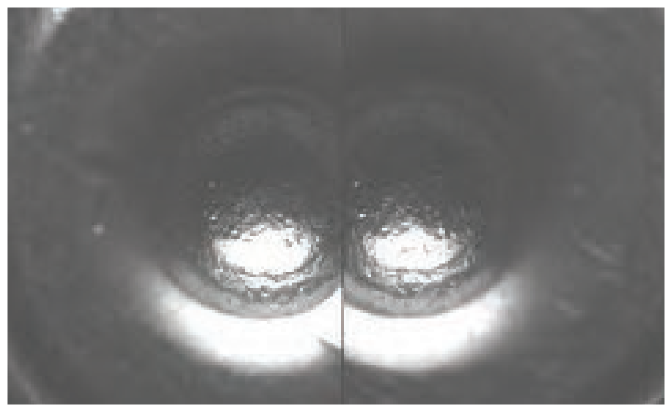


Figure 23: Comparison of firing pin impressions on test-fired cartridge case from Gun 1 (left) vs. Unknown Cartridge Case #2 (right), showing sufficient agreement for identification, 30X

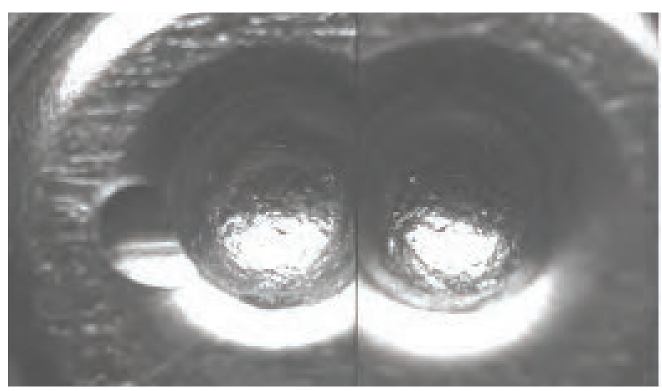


Figure 24: Comparison of firing pin impressions on test-fired cartridge case from Gun 3 (left) vs. Unknown Cartridge Case #2 (right), showing insufficient agreement for identification, 30X

References

[1] Berg, S. "The Drama of Forensic Ballistics," AFTE Journal, 1979, 11(3), pp. 44-48.
[2] Nichols, R., "The Scientific Foundations of Firearms and Tool Mark Identification-A Response to Recent Challenges," The CACNews, Second Quarter 2006, pp 8-27.
[3] Schwartz, A., "Challenging Firearms and Toolmark Identification—Part I," The Champion, October 2008, pp 10-19.
[4] "Understand the Causes: Unvalidated or Improper Forensic Science." The Innocence Project [Internet]. Available from: <http://www.innocenceproject.org>. Accessed 2011 Aug 8.
[5] Clarke, M. "Crime Labs in Crisis: Shoddy Forensics Used to Secure Convictions." Prison Legal News, 2010, 21(10), pp 1-22.
[6] Mnookin, J.L.; Cole, S.A.; Dror, I.E.; Fisher, B.A.; Houck, M.M.; Inman, K.; Kave, D.H.; Koehler, J.J.; Langenburg, G.; Risinger, D.M.; Rudin, N.; Siegel, J.; and Stoney, D.A., "The Need for a Research Culture in the Forensic Sciences," UCLA Law Review, 2011, 58(3), pp 725-779.
[7] National Academy of Sciences. "Strengthening Forensic Science in the United States: A Path Forward." Washington, D.C.: The National Academies Press, 2009.
[8] "FAQs About Clinical Studies". 2010 Mar 1. National Institutes of Health [Internet]. Available from: <http://www.cc nih.gov/participate/faqaboutcs.shtml>. Accessed 2011 Aug 8.
[9] Saks, M.J and Koehler, J.J. "The Individualization Fallacy in Forensic Science Evidence," Vanderbilt Law Review, 2008, 61(1), p. 201, footnote 9.
[10] Peterson, J.L.; Lin, G.; Ho, M.; Chen, Yingyu, and Gaensslen, R.E. "The Feasibility of External Blind Proficiency Testing. I. Background and Findings," Journal of Forensic Sciences, 2003, 48 (1), p. 6.
[11] National Academy of Sciences, *op. cit*., p. 207.
[12] "FAQs About Clinical Studies", *op. cit*.
[13] "CTS Statement on the use of Proficiency Testing Data for Error Rate Determinations." 2010 March 30. Collaborative Testing Services, Inc. (CTS) Forensics [Internet]. Available from: <http://www.ctsforensics.com/>. Accessed 2011 August 10.
[14] *Ibid*.
[15] Grzybowski, R.; Miller, J.; Moran, B.; Murdock, J.; Nichols, R.; and Thompson, R. "Firearm/Toolmark Identification: Passing the Reliability Test Under Federal and State Evidentiary Standards," AFTE Journal, 2003, 35(2), p. 216.
[16] Kaye, D.H; Bernstein, D.E.; and Mnookin, J.L. The New Wigmore: A Treatise on Evidence: Expert Evidence. R.D. Friedman, General Editor. Aspen Publishers (2nd ed. 2011), § 13.3.3, pp. 605-606.
[17] Expert Working Group on Human Factors in Latent Print Analysis. Latent Print Examination and Human Factors: Improving the Practice through a Systems Approach. U.S. Department of Commerce, National Institute of Standards and Technology, 2012, p. 33.
[18] DeFrance, C. and Van Arsdale, M.D. "Validation Study of Electrochemical Rifling," *AFTE Journal,* 2003, 35(1), pp. 35-37.
[19] Clow, C. "Cartilage Stabbing with Consecutively Manufactured Knives: A Response to Ramirez v. State of Florida," AFTE Journal, 2005, 37(2), pp. 86-116.
[20] Lightstone, L. "The Potential for and Persistence of Subclass Characteristics on the Breech Faces of SW40VE Smith & Wesson Sigma Pistols," AFTE Journal, 2010, 42(4), pp. 308-322.
[21] Smith, E.D. "Cartridge Case and Bullet Comparison Validation Study with Firearms Submitted in Case Work," AFTE Journal, 2005, 37(2), pp. 130-135.
[22] Peterson, et al., *op. cit*., pp. 7-10.
[23] Bunch, S.G., and Murphy, D.P. "A Comprehensive Validity Study for the Forensic Examination of Cartridge Cases," AFTE Journal, 2003, 35(2), pp. 201-203.
[24] Giroux, B.N. "Empirical and Validation Study: Consecutively Manufactured Screwdrivers," AFTE Journal, 2009, 41(2), pp. 153-158.
[25] Bunch and Murphy, *op. cit.*, p. 201.
[26] Giroux, *op. cit*., p. 156.
[27] Kaye, D.H. "Identification, Individualization, Uniqueness," Law, Probability & Risk, 2009, 8(2), p, 87, footnote 18.
[28] Schwartz, *op. cit.*
[29] "Understand the Causes: Unvalidated or Improper Forensic Science," *op. cit.*
[30] Clarke, *op. cit.*
[31] Mnookin, et al., *op. cit.*
[32] Diamond, B.L. "The Scientific Method and the Law." The Hastings Law Journal, Vol. 19, Nov. 1967, p.193.
[33] Grzybowski, R. and Murdock, J. "Firearm and Toolmark Identification - Meeting the Daubert Challenge," AFTE Journal, 1998, 30(1), pp. 6-8.
[34] Mnookin, et al., *op. cit.*, pp. 743-744.
[35] Supica, J. and Nahas, R., Standard Catalog of Smith & Wesson, Second Edition, WI: Krause Publications, 2001, p. 237.
[36] Murphy, Doug. "CTS Error Rates, 1992-2005 Firearms/ Toolmarks" [PDF document]. Originally presented at the AFTE Training Seminar in Henderson, Nevada on May 5, 2010. Retrieved from Scientific Working Group for Firearms and Toolmarks website: http://www.swggun.org/resources/docs/CTSErrorRates.pdf. Accessed June 10, 2013.
[37] Lightstone, *op. cit.*, p. 310-313.
[38] *Ibid.*, p. 312.

[39] Thompson, E. "9mm Smith & Wesson Ejectors," AFTE Journal, 2002, 34(4), pp. 406-407.

[40] Kennington, R.H. "Grinding Marks: Paradigm Lost." Unpublished paper presented at AFTE Annual Training Seminar in St. Louis, Missouri, June 2000. Kennington's presentation abstract, containing his experimental results and photographs, were provided to the author courtesy of John E. Murdock.

[41] "Range of Conclusions Possible When Comparing Toolmarks." AFTE Glossary, 5th edition, ver. 5.070207, pp. 118-119.

[42] Nichols, R.G. "Firearm and Tool Mark Identification: The Scientific Reliability and Validity of the AFTE Theory of Identification Discussed Within the Framework of a Study of Ten Consecutively Manufactured Extractors," AFTE Journal, 2004, 36(1), pp. 67-88.

[43] Biasotti, A., Murdock, J., and Moran, B. Chapter 35 "Firearms and Toolmark Identification", pp. 641-730 in Vol. 4, Modern Scientific Evidence: The Law and Science of Expert Testimony, 2011-2012 Edition (Faigman, DL, Blumenthal, JA, Sanders, J, Chen, EK, Mnookin, JK, and Murphy, EE.), Eagan, MN: Thompson-Reuters/West, p. 690.

[44] Ayoob, Massad, "The Glock 22: America's Best-Selling Police Pistol," Guns, July 2008, pp. 14.

[45] AFTE Criteria for Identification Committee Report: "Theory of Identification, Range of Striae Comparison Reports and Modified Glossary Definitions - an AFTE Criteria for Identification Committee Report", AFTE Journal, Vol. 24, No. 3, July 1992, pp. 336-340.

THE IDENTIFICATION OF CONSECUTIVELY RIFLED GUN BARRELS

By David J. Brundage, I-MC Forensic Services Agency

KEY WORDS: firearms identification, consecutively rifled, fired bullets, semi-automatic pistol.

ABSTRACT

Ten consecutively rifled Ruger P-85 pistol barrels were test fired to obtain standards and unknowns for comparison by a group of thirty firearm examiners from nationally accredited forensic laboratories. The manufacturing process was monitored to assure true consecutiveness. Each examiner was supplied a group of fifteen unknowns. At least one bullet was fired from each barrel, but some barrels could have had two, three, or four. The results showed that correct associations were made in all cases and that consecutive gun barrels could be properly identified.

INTRODUCTION

In 1991, there were over three million firearms manufactured in the U.S.. Unfortunately, some of them were used in the commission of assorted crimes. In 1992 alone, 15,377 murders were committed with a firearm [1]. Twenty-one percent of the guns submitted to the Chicago Police Department in 1993 involved a microscopic comparison of bullets or cartridge cases recovered from a crime scene.

The ability of the forensic firearm examiner to identify fired bullets to a specific firearm has been substantiated and is currently a court qualified practice. Conducting this type of examination has been the basis of previous work [2, 3, 4] where individual characteristics on the lands and grooves in a gun barrel have been shown to be unique to one particular firearm.

Several studies have supported the contention of uniqueness in multiple bullets or cartridge cases fired from one firearm. Kirby [5] identified all of the cartridge cases and 30 of the bullets in one of the earliest studies involving 900 tests. Ogihara, and others [6] identified a series of 5,000 bullets and cartridge cases fired from a single military 45 ACP pistol. Matty [7] found dramatic changes in the first few bullets fired from newly manufactured barrels, but subsequent firings "settled in" these characteristics and reproduction between tests stabilized.

Other studies have specifically examined consecutiveness. Lutz [8], using two gun barrels cut from one rifled barrel, fired three different types of bullets and had no difficulty identifying the correct barrel. Hall [9] utilized four barrels, two of which were consecutively rifled while the other two were randomly taken from general production. The author reported no difficulty in identifying the bullets fired from them, even though the barrels had been lapped (smoothed) after rifling. Murdock [10] examined consecutive

(Continued on page 439)

Authors note: The following research was performed while serving as the Firearms Training Coordinator for the Training and Applications Laboratory of the Illinois State Police Forensic Science Command, Carbondale Laboratory. The original paper was accepted as a master's thesis in 1994 at Southern Illinois University and has been presented at the 1994 AFTE Seminar in Indianapolis and again in 1995 at the American Academy's meeting in Seattle.

(Continued from page 438)
barrels from three different manufacturers and, after the first few tests, was able to identify each of them in their respective order.

Although these previous studies dealt with certain aspects of concern to the firearm examiner, some other important variables were omitted or, not addressed. The purpose of this study was to investigate whether two or more bullets, fired from consecutively rifled (manufactured) gun barrels, could be correctly associated with the barrel from which they were fired.

RESEARCH QUESTIONS

The following research questions were developed for project study:

1. Can trained forensic firearm examiners distinguish between two or more multiple gun barrels that have been consecutively rifled?
2. Can firearm examiners in the Illinois State Police Forensic Science Command differentiate individual characteristics of bullets fired from gun barrels that have been rifled in a consecutive manner?
3. Can firearm examiners from nationally accredited forensic laboratories accurately identify bullets that were fired from gun barrels that have been consecutively rifled?

Current conclusions made about the origin of forensic samples are based on the fact that bullets and cartridge cases can be positively identified to the gun from which they were fired. Extensive research has been published to support this hypothesis [3, 4, 11, 12].

Firearms are typically manufactured in production-line style. Each piece is fabricated utilizing several processes, sub-assembled, and then brought together with other pieces for final assembly. One of the tools used to make the spiral grooves in a gun's barrel, called a gang broach, has numerous sharp cutting edges which wear down during use. Imperfections on these cutting surfaces cause corresponding imperfections on the surfaces of the lands and grooves (rifling) in the barrel. These imperfections occur throughout the length of the barrel and make each gun barrel unique, no matter what rifling method is used [4, 11, 13, 14]. Bullets fired through these rifled barrels are marked by both the rifling and these minute imperfections. These "toolmarks", in the form of scratches (called striae), can be observed and identified with other bullets fired through the same barrel using a comparison microscope.

In a 1987 lecture given to a Forensic Science, Civil and Criminal Symposium held in Eugene, Oregon [15], a retired crime laboratory director stated that firearm examiners could not conclusively identify consecutively rifled gun barrels. Such accusations, if true, would terminate the use of firearms evidence as conclusive information in criminal investigations, court proceedings, and ultimately, the training of new examiners.

Although previous research gives some credence to identifying consecutively rifled gun barrels, variables have limited their credibility. In most studies the examiners knew which barrels fired the test bullets, and could not verify true consecutiveness of manufacturer, or each barrel's orientation with respect to rifling.

This study established the examiner's ability to correctly associate consecutively manufactured gun barrels and eliminate the deficiencies of previous research. Statements about general identifiability of bullets were also proven incorrect. The results of this study have provided the forensic community with additional supportive documentation for all firearm identifications.

EXPERIMENTAL

A nine millimeter model P-85 semiautomatic pistol and ten consecutively manufactured barrels were obtained from the Sturm, Ruger Company. The nine millimeter caliber was chosen based upon the current popularity of the cartridge and the manufacturer's willingness to provide the firearm and barrels for the project.

(Continued on page 440)

(Continued from page 439)

To observe the manufacturing process and verify consecutiveness of the barrels, travel was arranged to the Ruger factory in Prescott, Arizona. Ten heat-treated, rough cast barrels were die stamped "11 through 20" to establish the order for final broaching, rifling, and finishing. Each completed barrel was checked for proper head space using the same frame and slide, and then test fired with fully loaded magazines at the factory.

Two lots of ammunition were obtained for this project. One was supplied by the Winchester/ Olin Corporation and the other was purchased from a local source. To provide the standards for each test set, two bullets were fired from each gun barrel. These were marked with the same number of the barrel plus an "A" or a "B" to distinguish them from each other and from the other nine barrels. The test standards (TS) from each barrel were packaged together, but separated from other TS bullets.

The questioned (Q) bullets were fired from the same ammunition previously listed. Each was marked from "1" through "1200" and placed in a single container. No special selection of the Q bullets was attempted when drawn from the container for firing in any of the barrels. A relatively close association was maintained between firing of the two TS bullets and the Q bullets in each particular test set to assure they were the "best evidence" [4].

Each group of Q bullets of those marked above were randomly chosen and fired in one of the gun barrels. These were placed in a container and selected later for assembly of the final set of Q bullets. This was repeated for each of the ten barrels and for each of the thirty sets made for the sample population.

The final group of Q bullets for each test set consisted of 15 bullets randomly drawn from each of the ten groups fired from each of the barrels. An independent assembler was used to select the 15 Q bullets. After selection each Q bullet was placed in an individual coin envelope to keep them separate from the other Q and TS bullets. There was at least one Q bullet selected from each barrel for every test-set, which supplied ten of the 15 Q bullets. The other five Q bullets were selected from any combination of those remaining.

Therefore, each test set provided to the assembler for selection consisted of a total of 60 bullets; 20 test standards (TS) and 40 questioned (Q) bullets, a combination from from each barrel. The assembler then selected 15 Q bullets from the group of 40 and packaged them together with the 20 TS bullets previously packaged. This was repeated until all thirty test sets were assembled. Each test-set contained a total of thirty-five bullets.

Each completed test set was sealed in a manilla envelope and packaged in a padded mailer for shipment. A key, indicating which Q bullets were selected for each test set, was returned to this researcher after all thirty test sets were sealed. The thirty firearm examiners used in this sample population were chosen from laboratories that were nationally accredited by the American Society of Crime Laboratory Directors - Laboratory Accreditation Board (ASCLD-LAB).

RESULTS AND DISCUSSION

There were no incorrect answers for any of the results collected. Each examiner made the proper associations between each gun barrel and all Q (unknown) bullets from each of the thirty test sets. Table 1 shows all ten gun barrels with their corresponding number of bullets of each test set identified in the survey. The average time required to conduct the examination and compare each test set was approximately nine hours.

One laboratory did not have an answer for one of the barrels, but also had one bullet that was not identified to any of the barrels. It was stated that none of the bullets could be positively identified with that particular barrel, nor could any of the barrels be identified with that particular bullet. This is an "inconclusive" answer. Inconclusive responses were not considered incorrect

(Continued on page 441)

TABLE #1
BULLETS IDENTIFIED FOR EACH BARREL

TEST SET NUMBER	1	2	3	4	5	6	7	8	9	10
1	1	1	1	1	1	1	1	4	2	2
2	2	2	1	1	1	1	1	1	2	3
3	1	1	2	1	1	1	1	3	2	2
4	2	1	1	1	1	1	1	2	3	2
5	1	1	1	1	1	1	1	3	2	3
6	1	1	3	1	1	1	1	2	3	1
7	2	2	1	1	1	1	1	1	3	2
8	1	1	1	1	2	1	1	3	1	3
9	1	1	1	1	1	1	1	4	3	1
10	2	1	1	1	1	1	1	2	2	3
11	1	1	1	1	1	1	1	3	3	2
12	1	1	1	1	1	1	1	3	3	2
13	1	1	1	1	1	1	1	4	2	2
14	1	1	1	1	1	1	1	2	3	3
15	1	1	1	1	1	1	1	3	2	3
16	1	2	1	1	2	1	1	1	3	2
17	1	1	2	1	1	1	1	3	2	2
18	1	1	1	1	1	1	1	4	3	1
19	2	2	1	1	1	1	1	2	1	3
20	1	1	1	1	1	1	2	3	2	2
21	1	1	1	1	1	1	2	2	3	1
22	1	2	1	1	1	1	1	2	1	4
23	1	1	1	1	1	1	1	4	3	1
24	1	1	1	1	1	1	1	2	4	2
25	3	1	1	1	1	1	1	3	1	3
26	1	2	1	1	1	1	1	2	3	2
27	1	1	1	1	1	1	2	3	3	1
28	1	1	1	1	1	2	1	2	1	4
29	2	1	1	1	1	1	1	3	2	2
30	1	1	1	1	1	1	1	4	2	2

GUN BARREL NUMBER

(Continued from page 440)

because there are many factors that can affect identifiability. These factors include poor markings (striae) on either test or evidence, a lack of reproduction of these markings on the tests, or even a lack of correspondence of the individual characteristics between both test and evidence. To maintain the degree of realism to normal casework in this research design such a response is also not considered a wrong answer. For this specific test set there was only one bullet that could be identified to that particular barrel, and the bullet not identified was the correct answer.

A pre-test was administered to five accredited laboratories which did not participate in the final test. The test sets were assembled as described above and mailed to a firearm examiner in each of the chosen laboratories. Three specific areas of the results were examined in the pre-test including construction, assembly, and answer sheet.

(Continued on page 442)

(Continued from page 441)

All five pre-tests were completed on schedule. Their results suggested adjustments to the packaging, construction, and marking of the bullets. Bullets for the master test were individually packaged to prevent possible surface damage and additional care was exercised in marking them. The test results and remaining comments supported both the construction of the test and the viability of the questionnaire.

A discrepancy in the pretest was discovered by each of the pre-test examiners when it was determined that one of the barrels did not get changed during the firing sequence. This provided two sets of Q bullets for one of the barrels and none for the next. The bullets were marked as if these barrels had been properly changed and test fired. This discrepancy was either noted on the questionnaire or by direct contact with the researcher.

Background information provided through these results gave insight about firearm examiners of nationally accredited laboratories. For example, the average number of years of experience in this field was 18 years, while the amount of time spent in training was almost two years. Only six of the 30 examiners surveyed participated in a formal training program, while 19 were trained "on-the-job". Other areas queried by the survey are listed below:

Employment

--City or County Agency	5
--State Agency	25
--Sworn Officers	7
--Non-sworn	23

Microscopes

--Leica	20
--Leitz	7
--Other	3

Lighting

--Fiber Optic	14
--Fluorescent	9
--Incandescent	7

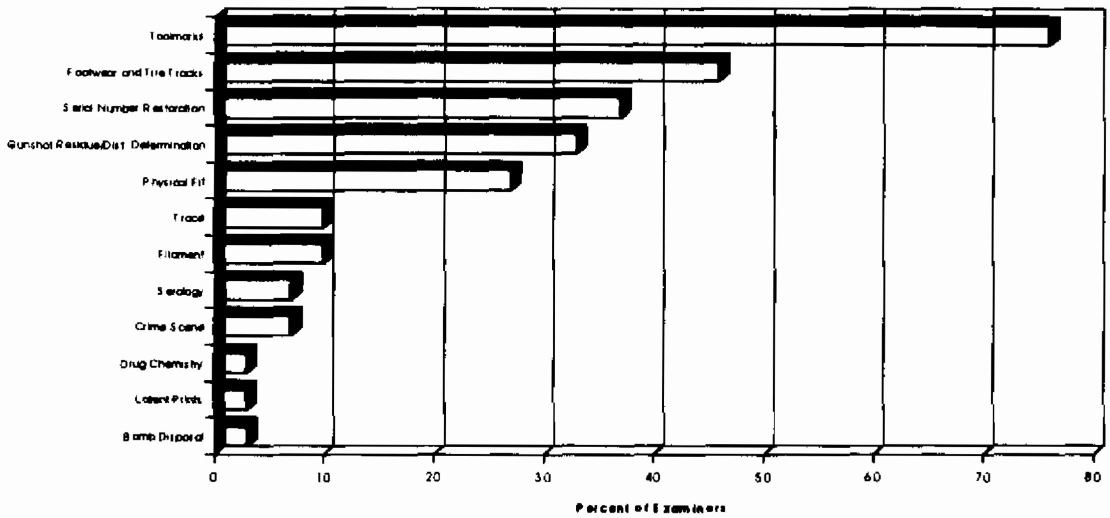
Twenty-eight examiners said they routinely examined other types of evidence besides those involving firearms.

The type of evidence examined and percentage of examiners conducting these examinations cover a broad array of analyses. While toolmark identification is the largest type, it does not include all examiners. Areas like Serology and Drug Chemistry are not unusually combined with firearms identification. The bar graph in Table 2 illustrates this data.

All but one of the examiners belonged to a professional forensic organization. The most prevalent was the Association of Firearm and Tool Mark Examiners (AFTE). The specific organization and percentage of membership follow in descending order of membership. Some individuals belong to more than one organization. Examiners in any particular field seem to identify with certain forensic associations. For example, AFTE is the only organization that is specifically for examiners conducting firearm and toolmark identification work. Most of the others cover several areas including firearm identification. AFTE has published a formal training outline (manual), Glossary of terms specific to the field, and a Quality Assurance Program. Others, like California Association of Criminalists, have established certification programs for its members. AFTE is currently setting up such a program, but it is not in place as of this writing.

(Continued on page 443)

Table #2 - Examiners Conducting Other Examinations



(Continued from page 442)

Association of Firearm and Tool Mark Examiners (AFTE) 97%

Midwestern Association of Forensic Scientists (MAFS) 33%

International Association for Identification (IAI) 27%
(National Chapter)

Northwest Association of Forensic Scientists (NWAFS) 10%

Southwest Association of Forensic Scientists (SWAFS) 10%

International Association for Identification (IAI) 10%
(Illinois, Missouri, and Texas Chapters)

California Association of Criminalists (CAC) 3%

Southern Association of Forensic Scientists (SAFS) 3%

International Association of Bomb Technicians and Investigation (IABT&I)

American Society of Crime Laboratory Directors (ASCLD) 3%

It should be noted that, in addition to the examiners who were official participants in this study, an additional fifteen to thirty examiners participated unofficially and each obtained the same correct answers.

CONCLUSIONS

This project verifies the three research questions regarding the identification of consecutively rifled gun barrels. The results also demonstrate that, on a national level, properly trained examiners can distinguish two or more bullets fired from such barrels. Furthermore, they can accurately differentiate the individual characteristics of test shots fired from consecutively rifled barrels. This data also shows that not only are consecutively rifled gun barrels different from each other, but are unique and can be differentiated.

(Continued on page 444)

(Continued from page 443)

ACKNOWLEDGEMENTS

I wish to express my gratitude to the Sturm Ruger Company for providing the firearm and barrels that were used in this project, as well as the Winchester/Olin Corporation for supplying the ammunition. I also want to thank Dr. Arthur L. Casebeer of Southern Illinois University at Carbondale for his guidance in assembling this research, the Illinois State Police Forensic Services Command and Carbondale Laboratory for the support necessary to conduct this project, and the Indianapolis-Marion County Forensic Services Agency for the time to write this article for publication. Finally, I want to give a special acknowledgement to the examiners that participated in this project; whose time and effort actually made this project a success, not only to the field of Firearms Identification, but also for firearm examiners, past and future, who may have to deal with negative comments about identifiability on the witness stand.

REFERENCES

[1] FBI, *Crime in the United States: Uniform Crime Reports,* Washington DC: US Department of Justice, 1992.

[2] Goddard, C. H., "A History of Firearms Identification," *Chicago Police Journal,* Chicago, 1936, pp. 1-15.

[3] Davis, J. E., *An Introduction to Tool Marks, Firearms, and the Striagraph,* Charles C. Thomas, Springfield, 1958.

[4] Mathews, J. H., *Firearms Identification,* (2nd. ed.), Charles C. Thomas, Springfield, 1973.

[5] Kirby, S. J., "Comparison Of 900 Consecutively Fired Bullets And Cartridge Cases From A 455 Caliber S & W Revolver," *A.F.T.E. Journal,* Vol. 15, No. 3, 1983, pp. 113-126.

[6] Ogihara, Y., Kubotga, M., Sanada, M., Fukuda, K., Uchiyama, T., and Hamby, J. E., "Comparison Of 5000 Consecutively Fired Bullets And Cartridge Cases From A 45 Caliber M1911A1 Pistol," *A.F.T.E. Journal,* Vol. 15, No. 30, 1983, pp. 127-140.

[7] Matty, W. A., "Comparison Of Three Individual Barrels Produced From One Button Rifled Barrel Blank," *A.F.T.E. Journal,* Vol. 17, No. 30, 1985, pp. 64-69.

[8] Lutz, M. C., "Consecutive Revolver Barrels," *A.F.T.E. Newsletter,* Vol. 2, No. 9, 1970, pp. 24-28.

[9] Hall, E. E., "Bullet Markings From Consecutively Rifled Shilen DGA Barrels," *A.F.T.E. Journal,* Vol. 15, No. 1, 1983, pp. 33-53.

[10] Murdock, J. E., "A General Discussion of Gun Barrel Individuality And An Empirical Assessment Of The Individuality of Consecutively Button Rifled 22 Caliber Rifles," *A.F.T.E. Journal,* Vol. 17, No. 3, 1981, pp. 64-69.

[11] Hatcher, J. S., *Textbook of Firearms Investigation Identification and Evidence,* Telegraph Press, Harrisburg, 1935.

[12] *Firearms Identification,* Institute of Applied Science, Chicago, 1954.

[13] Howe, J. V., *The Modern Gunsmith, Two Volumes in One,* Bonanza Books, New York, 1982.

[14] Wallack, L. R., *Encyclopedia of American Gun Design and Performance,* Winchester Press, 1983.

[15] Fox, R. J., *The Myth of Bullet Matching,* Forensic Science, Civil and Criminal Symposium conducted in Eugene, Oregon, November, 1987.

The Identification of Bullets Fired from 10 Consecutively Rifled 9mm Ruger Pistol Barrels: A Research Project Involving 507 Participants from 20 Countries

By: James E. Hamby, Ph.D., International Forensic Science Laboratory & Training Centre, Indianapolis, IN, David J. Brundage, M.S., Independent Examiner, Nashville, TN and James W. Thorpe, Ph.D., Senior Lecturer (Retired), University of Strathclyde, Glasgow, Scotland

Key Words: Automated Land Identification System (ALIS), BulletTRAX-3D®, consecutively rifled barrels, criteria for identification, Daubert, firearms identification, fired bullets, SciClops®, scientific research

ABSTRACT

Ten consecutively rifled RUGER P-85 pistol barrels were obtained from the manufacturer and then test fired to produce known test bullets and 'unknown' bullets for comparison by firearms examiners from around the world. This study is a continuation of one originally designed and reported on by David Brundage [1]. The original study was primarily limited to examiners from nationally accredited laboratories in the United States. For this study, the sets were provided to firearms examiners around the world. The RUGER P-85 pistol and the 10 consecutively rifled barrels used for the original study were borrowed from the Illinois State Police. Ammunition was obtained from the Winchester Ammunition Company (A Division of Olin) and 240 tests sets were produced and distributed to forensic scientists and researchers worldwide. A thesis which involved a total of 201 participants – including the original 67 reported on by Brundage – was published by Hamby and Thorpe in 2001 [2]. This paper reports the final conclusions of the research conducted by Brundage, Hamby and Thorpe over a 10 year period [3, 4].

Introduction

Current practices in firearm and toolmark identification training and actual laboratory casework are based on the theory that fired bullets and fired cartridge cases can be identified to the firearm that fired them. A forensic scientist trained in firearm and toolmark identification is often able to specifically identify, or eliminate, a firearm involved in a shooting when that firearm is evaluated in conjunction with recovered evidence. Extensive research has been conducted and published by forensic firearm and toolmark examiners during the past 100+ years to support this theory .

A firearm and toolmark examiner microscopically evaluates fired ammunition components using an optical comparison microscope. For fired bullets specifically, the fine scratches (striae) found on the bearing surfaces are assessed. These striations are considered to be accidental in nature and to arise from randomly occurring imperfections during the manufacture of the gun barrel. Because these imperfections occur at random, the pattern of striations is considered to be unique to a common origin, such as a specific firearm or tool. In the case of a fired bullet, the striations are impressed on the bullet by force and motion as the bullet travels down the barrel of the firearm. Although this specific research project pertains to bullets fired from consecutively manufactured barrels, the same type of analytical techniques and laboratory equipment are used when examining fired bullets, fired cartridge cases and a wide variety of tools – whether from different firearms and tools and/or consecutively manufactured firearms and tools.

Numerous studies have shown that a properly trained firearm and toolmark examiner has the ability to identify fired bullets and fired cartridge cases to the firearm that fired them, even when multiple bullets and cartridge cases have been successively fired. Another area of concern is the examination of bullets and cartridge cases fired by different firearms.

Concerning firearm barrels, it is recognized that striations are caused by imperfections in the rifling tools during the barrel manufacturing process and also can be inherent in the manufacturing process itself. The rifling tools wear during their use and potentially impart a continually changing set of striations, and the machining process itself does not yield identical barrels on the microscopic level. It would be expected that the greatest potential for similarity of striations would be encountered with firearm barrels that are consecutively rifled using the same rifling tool.

Date Received: September 15, 2008
Peer Review Completed: March 28, 2009

Reproducibility of Striae and Impressed Marks (Consecutive and Non-Consecutive)

The following research articles – listed in chronological order – reflect a very small number of the overall research that has been conducted involving consecutively manufactured components as well as randomly (non-consecutive) manufactured components (barrels, extractors, ejectors, breech faces, knives, etc.).

One of the first recorded identifications of a specific fired projectile to a firearm occurred in 1898 in Neuruppin, Germany. Professor Paul Jeserich, a gifted forensic chemist from Berlin, was requested by the Neuruppin district court to compare a bullet removed from the body of a murder victim to a revolver owned by a suspect [7]. Jeserich test fired the revolver and then carefully produced a series of photomicrographs of the murder bullet and the test fired bullet. When he compared the photographs, he observed abnormalities on the bullets that indicated that both had been fired from the same firearm. His testimony was instrumental in the conviction of the defendant. His other interests, however, precluded his continuing further research into the area of firearm identification.

In the United States in 1907, the first recorded examination of multiple firearms in conjunction with fired cartridge cases involved inspectors at the US Army's Frankford Arsenal. The arsenal staff examined 279 service rifles and 33 fired cartridge cases from a shooting incident. The rifles were test fired and the test cartridge cases examined in conjunction with the evidence cartridge cases. The staff reported that they were able to identify some of the cartridge cases to the rifles. Their conclusions are an excellent example of early cartridge case identification [8, 9, 10].

Additional research continued in this forensic field during the next 25 years by early self-trained examiners such as Sydney Smith, Robert Churchill, Dr. Calvin Goddard and others. Four heavily reported criminal events permanently established the discipline of firearm and toolmark identification in both the United Kingdom and the United States. These cases involved the assassination of the Sidar in Egypt, the murder of Constable Gutteridge in England, and the Sacco-Vanzetti murder case and St. Valentine's Day Massacre in the United States [11, 12, 13]. The ability of these pioneer examiners to identify both fired bullets and fired cartridge cases to a specific firearm was instrumental in establishing firearm and toolmark identification as one of the forensic sciences.

Numerous studies support the contention of uniqueness where multiple bullets and/or cartridge cases are fired from one firearm. An excellent article by Bonfanti and De Kinder [14], discusses several scientific studies (some of which are mentioned in this article) that have been conducted where fired bullets and/or cartridge cases have been examined after test firings from consecutively manufactured firearms. In other instances, research has been conducted to evaluate fired components from a large number of firearms.

Two excellent articles and a presentation by Ronald Nichols [15, 16, 17] comprise a comprehensive review of the literature that pertains to firearm and toolmark identification criteria. Additional articles such as ones by Grzybowski, Murdock, Moran, Biasotti and others [18, 19, 20] offer a valuable compendium of reference materials that discuss scientific methods, reliability and the validity of the field of firearm and toolmark identification.

Numerous historical articles have been published [21, 22, 23, 24] which also provide various references concerning the field of firearm and toolmark identification. Other researchers such as Biasotti, Murdock, Moran, Thompson and many others have conducted extensive research and published their findings [25, 26, 27, 28, 29, 30, 31]. Due to space limitations, and the nature of this specific research project, only a few references are provided below.

In 1930, a rod of steel (barrel blank) was bored and rifled at an U.S. Government arsenal. The barrel stock was rifled and then cut into six pieces to form six short barrels. A bullet was test fired from each of the six barrels and scribed with a secret marking. Colonel Goddard was given the six scribed bullets and six barrels for evaluation and examination. In this blind study, Goddard correctly associated the scribed bullets to the appropriate barrel [32].

In 1957, Flynn reported on a study in which the Chicago Police Department (CPD) Crime Lab examined a total of 100 consecutively manufactured chisels that had been finished using a grinding process. He reported that a total of 5,050 total comparisons were made during the experiment with no misidentifications [33].

In 1958, Kirby fired 900 lead bullets from a .455 caliber revolver and was able to identify that all of the cartridge cases had been fired in the same weapon [34]. However, he was only able to identify the first thirty bullets as being fired from the revolver because the patterns of striations on the bullets were affected by the barrel becoming leaded during the test.

In 1970, Lutz used two consecutively rifled and machined revolver barrels for a 38 Special caliber Smith & Wesson Model 10 revolver. Three different types of bullet configurations, including lead bullets, were test fired and examined. Of those participating in the examination of the test fired bullets, none had difficulty differentiating between the proper barrels [35].

In a study conducted in 1972, a total of 501 full metal jacket (FMJ) projectiles were fired from an M16A1 223 caliber assault rifle [36]. The assault rifle was selected from the Laboratory Weapons Reference Library (WRF) while the ammunition used was from the Laboratory's Ammunition Reference File (ARF) [32, 33]. The 501 cartridges were fired – using the full automatic mode - as fast as the 20 round magazines could be changed and every hundredth projectile collected in a cotton recovery box. It was possible to microscopically identify all the bullets as having been fired by the same rifle.

In 1972, Murdock compared bullets fired from four crowned, button-rifled barrels with bullets fired from the same barrels after they had been recrowned. Although he observed some changes in the rifling, he could still associate the proper bullet to the specific barrel. Another set of test fired bullets was compared to the first set after the barrels were recrowned a second time with a similar result. This study demonstrated that the crowning process had minimum effect on identifying fired bullets [37].

In 1973, an U. S. Army Captain was shot and killed while standing in his tent in a bivouac [encampment] area. The assailant fired a 223 (5.56mm) caliber M16A1 assault rifle at the Captain's shadow in the tent. Investigators seized a total of 47 M16A1 assault rifles from personnel in the bivouac area. The rifles, along with the fired bullet components recovered during the autopsy, were forwarded to the US Army Criminal Investigation Laboratory at Fort Gordon, Georgia. Special Agent John G. Ward, Sr., senior firearms examiner for the laboratory, test fired the 47 rifles and microscopically compared the test-fired bullets to the evidence bullet fragments. Ward was able to identify the rifle used to shoot and kill the Captain. The suspect, a disgruntled soldier, was found guilty of murder [38].

Butcher & Pugh reported on a study in 1975 which involved the examination of test marks made by ten consecutively made bolt cutters as well as ten randomly selected bolt cutters – all with ground working surfaces (blades). The study showed no more than 29% matching striae for known non-matches (KNM) and between 87% and 93% matching striae for known matches (KM). The implication of this research suggests that there is no risk of misidentification by a competent examiner [39].

Ogihara, and others, conducted an extensive research study in 1977, by examining 5000 bullets and cartridge cases fired by an U.S. Army issue M1911A1, 45 (11.45mm) ACP caliber semiautomatic pistol [40]. The researchers used standard 45ACP caliber FMJ military ammunition for the project and collected every tenth fired bullet and cartridge case for examination. The firearm used for the project was part of the National Research of Police Science Institute's (NRIPS) Weapons Reference Library and the ammunition was provided by the U.S. Army Criminal Investigation Laboratory – Pacific (USACIL-PAC) – now closed. This study involved firearm and toolmark examiners from three forensic laboratories and required a substantial amount of time to effect the comparisons for the bullets and cartridge cases. Using standard microscopic techniques, the researchers were able to identify all of the bullets and cartridge cases as having been fired by the same pistol.

In 1978, Watson published an article discussing the uniqueness of two consecutively manufactured knives. His research revealed that no carryover of individual markings was found to exist between the two knives and that the knives could be individually identified [41].

Cassidy reported on a study in 1980 where he examined the individuality of striated marks produced by consecutively broach cut tongue and groove pliers. His examination and observations of the jaw teeth and their test marks revealed no subclass marks and that the striated marks produced are individual to the tool that made them [42].

For a comprehensive study in 1981, Murdock obtained three consecutively button-rifled 22 caliber (5.56mm) barrels each from three different manufacturers. The nine barrels were machined to fit one bolt-action rifle. Thirty lead bullets were fired from each of the nine barrels and compared to each other. His research determined, as in other studies, that the first few bullets fired from each barrel were not identifiable to each other. The remaining bullets, from each barrel, were identifiable to each other and could be distinguished from those fired from the other barrels [43].

In 1982, Tuira compared two consecutively manufactured Buck brand knives that were used to cut inflated tires. His microscopic observations of the resulting toolmarks determined that the toolmarks were significantly different [44].

In a study by Hall in 1983, four barrels in 308 caliber (7.62mm) with polygonal rifling were used. Two of the barrels were consecutively rifled while the other two were randomly taken from the production line. Hall reports that he encountered no difficulties in identifying bullets fired in any of the barrels. He used three different brands of ammunition with the first five bullets fired from each barrel used for stabilizing the pattern of striations. The bullets, fired after the first five, were identifiable to each other and could be distinguished from those fired in any other barrel. Hall observed some change in striae when comparing bullets that were sequenced further apart from each other, but this did not preclude identification [45].

In 1983 Shem and Striupaitis fired 501 bullets and cartridge cases using a Raven Model P-25 25 (6.25mm) Auto caliber semiautomatic pistol. The researchers collected every 10th fired bullet and cartridge case for examination. They concluded that, although changes were occurring in the bullet striae and breechface marks, it was possible to identify bullet 1 to 501 as well as cartridge case 1 to 501 [46].

In 1984, Matty and Johnson examined the concentric marks produced by Smith & Wesson firing pins. Subclass characteristics were found and determined to be a result of the lathe mounted cutter being much harder than the firing pins. The researchers also determined that areas of the firing pins that contain random breaks in the striated lines due to metal tearing or areas that show wear can be used for identification [47].

Matty conducted a study in 1984 involving three consecutively made breechfaces from Raven semiautomatic pistols. His observations were that the concentric toolmarks on the breechfaces could be individualized and that the toolmarks were not subclass [48].

In 1985, Matty reported on a project involving the examination of three individual barrels produced from one button-rifled barrel blank. He noted some subclass characteristics in the groove impressions but not in the land impressions. He also determined that the striae changed significantly during the first few test firings [49].

In 1985, Van Disk reported on his examination of fifty steel marking stamps made from the same hob (die). The marking stamps were examined for subclass marks. Van Disk determined that unique defects from the hobbing process could be used to correctly identify each stamp [50].

Uchiyama conducted a study in 1986 where he examined the breechface marks produced by 25 Auto caliber Browning, Raven and Titan semiautomatic pistols. He determined that subclass characteristics were significant and informed examiners to be cautious when examining these types of firearms [51].

In 1986, Dr. Gross - then head of the Bundeskriminalamt (BKA) firearms section - reported on a high profile murder case that had occurred in Germany in 1984 & 1985. The case involved test firing some 7,862 similar type pistols with the test fired items submitted for examination. The examiners identified test fired components from pistol number 3,704 [52].

In 1992, Schecter and others test fired a 223 caliber (5.56mm) GALIL rifle 7,100 times, using a variety of 223 caliber ammunition. The researchers microscopically examined the fired cartridge cases specifically for the ejector marks because the ejector on a GALIL rifle is part of the rifle and is not removable. Schecter and others were able to identify the ejector marks on the casings with a spread in excess of 7,050 firings [53].

In 1992, Hall performed a series of tests in which consecutive test cuts in lead were made with bolt cutters. Hall reported that lead is a suitable material for test marks and that cuts in shackles may or may not change the tool depending upon the hardness of the shackle [54].

In 1994, Thompson reported on a follow-up study of the article by Matty on Raven breechfaces. He obtained four breechfaces from Phoenix pistols (formerly Raven) and compared them to determine the nature of their marks. His examination confirmed the findings of Matty that breechfaces possess unique identifying marks [55].

Brown & Bryant, in 1995, reported on a study of multi-barreled derringers in which it was assumed that the barrels were rifled consecutively. In one instance, one set of derringer test fires showed some good correspondence in the groove impressions (gross marks), but showed little correspondence in the land impressions [56].

In 1996, Thompson examined the manufacturing process of Lorcin pistol breechfaces. He noted that Lorcin breechfaces were produced by stamping and then painted over - as opposed to being machined - and that false identifications could be possible if the only marks considered were from the breechface [57].

In 1998, Tulleners and Guisto obtained a Thompson Center Contender button rifled barrel which was sectioned one inch at a time after each test firing. A total of six sections were removed from the barrel. The bullets test fired from each sectioned barrel were compared to each other to determine how much the Consecutive Matching Striation (CMS) count had changed. Striae on the bullets were found to be significantly altered from one barrel section to the next. The results obtained from adjacent barrel sections were apparently comparable to the results Biasotti had obtained from different, uncut barrels [58].

Tulleners and Hamiel reported on a study in 1999 where the potential for subclass characteristics in Smith & Wesson revolver barrels was discussed. The article points out that a firearm and toolmark examiner should be careful when examining the groove impressions on fired bullets from barrels that have been rifled using broach rifling techniques [59].

In 2000, Miller reported on a study where he pushed bullets through two consecutively broached 44 caliber barrels. He examined the test bullets using the Biasotti/Murdock conservative CMS criteria for identifications and reported that there were no misidentifications [60].

Rosati reported on a study in 2000 involving the examination of four bunters that were produced using Electrical Discharge Machining (EDM). The bunters were used by Remington for the manufacture of 45 Auto caliber cartridge cases. Rosati's examination confirmed the random nature of marks from the EDM process on headstamp characters [61].

In 2000, Lopez and Grew conducted a study involving firearm bolt faces machined with an end mill. The study warns that a misidentification is possible unless the identification is based on breechface wear or machining "chatter" marks on the breechface [62].

In 2001, Hamby reported on the microscopic examinations of four 9mm cartridge cases that were test fired in 617 Glock Model 17 & 19 semiautomatic pistols. Hamby microscopically examined the cartridge cases against each other to validate that uniqueness and individuality exist among the fired cartridge cases. The observations were that each casing could be identified to the specific firearm [63].

In February 2001, at the American Academy of Forensic Sciences Meeting in Seattle, Washington, Brett Doelling presented the results of research that he had conducted involving multiple bullets fired from the same firearm. Doelling test-fired 4,000 cartridges using a 9x18mm caliber Makarov semiautomatic pistol. Every 100th bullet was collected and examined microscopically. Doelling concluded that although the markings continued to change, the 4000th bullet was identifiable to the 1st bullet [64].

In 2001, Miller, using a test set containing bullets from the Hamby & Brundage Ruger ten barrel test, reported that he had identified some very minor subclass characteristics but not sufficient to cause a misidentification. He also applied the conservative CMS Criteria which resulted in no misidentifications [65].

Eckerman reported on a study in 2002 in which toolmarks made by consecutively manufactured and belt- sanded chisels were examined for the possibility of subclass marks. Eckerman's examinations revealed that the marks were found to be individual to each chisel [66].

Lee reported on a study in 2003 where she used five consecutively manufactured screwdrivers to test the reproducibility of marks produced at various angles with both pushing and pulling motions. The toolmarks from each of the screwdriver blades were found to be individual to tool that produced them [67].

In 2003, Thompson & Wyant visited a knife production facility where they observed the actual production of 10 consecutively manufactured knife blades. The researchers produced a number of test sets containing known and unknown knife cuts using those 10 consecutively manufactured knife blades. The test sets were provided to firearm and toolmark examiners for examination. This test – the Knife Identification Project (KIP test) – demonstrated the ability by the majority of participants to successfully differentiate toolmarks made by consecutively manufactured knife blades [68].

Bunch and Murphy reported in 2003 on a study in which 10 consecutively manufactured Glock semiautomatic pistol slides were obtained from the factory in Austria. The manufacturing process of the 10 slides - which contain the breechface - was observed and the slides then used to produce test fired cartridge cases for a comprehensive validity study by examiners in the FBI Laboratory's Firearms-Toolmark Unit (FTU). Using breechface marks, the examiners were able to correctly identify cartridge cases fired by each of the different slides [69].

Vinci, and others, conducted an extensive study in 2004 that involved 2500 cartridges fired by a 45 (11.45mm) ACP caliber Springfield Armory semiautomatic pistol. The researchers examined every 100th fired cartridge case to evaluate sequential changes in both class and individual characteristics and reported that it was possible to identify all 2500 cartridge cases as having been fired by the recently produced pistol [70].

In 2005, Clow reported on an extensive research study that utilized 10 consecutively manufactured knife blades in a stabbing motion to determine if the marks produced were unique, reproducible and identifiable in pig cartilage. The toolmarks were found to be unique to each knife blade, reproducible and potentially identifiable in cartilage [71].

Smith reported on a research study in 2005 that was designed to test the accuracy of examinations by trained firearm and toolmark examiners who use pattern recognition as a method for identification. Eight FBI examiners took the test which consisted of both bullets and cartridge cases. No false positives or false negatives were reported [72].

In 2005, Collins reported on an empirical study involving the uniqueness of impressed toolmarks. He used twenty worn hammers to produce a series of test toolmarks and examined the marks to determine if they could be considered unique. His conclusions were that the marks could be considered unique [73].

In 2008, Gouwe, Hamby and Norris reported on a experiment that involved a total of 10,000 fired 40 S&W caliber cartridge cases using a Glock Model 22 firearm. The researchers microscopically examined every 10th cartridge case and determined that sufficient individual markings were present on the fired cartridge cases to identify cartridge case 1 to cartridge case10,000 [74].

Experimental Design

During the past eighty years, a significant volume of research has involved the evaluation of test fired bullets and cartridge cases. The research cited in this paper has included test firing a firearm numerous times to evaluate changes in microscopic characteristics observed on the fired components and also the test firing of consecutively rifled firearms to determine if bullets could be identified to the barrel from which they were fired. In every research project involving the examination of consecutively manufactured tools – including bullets from consecutively rifled barrels – the results have established that properly trained firearm and toolmark examiners have the ability to identify toolmarked surfaces to the correct tool. Despite the wealth of research, there are still challenges to this type of evidence in the courts system.

Brundage's original research study was expanded to examine the ability of numerous firearm and toolmark examiners on a worldwide scale to associate bullets fired from consecutively manufactured gun barrels as well as to provide test sets for training use within the participant's own laboratory. (Originally, the 67 participants were comprised of 30 official examiners that were from ASCLD/LAB Accredited Laboratories, 30 unofficial examiners that were from non-accredited laboratories, and 7 examiners that were requested to conduct a pre-test evaluation of the test sets prior to distribution. At that time, all 67 participants were from laboratories in the United States)

This experiment was undertaken to address some of the following issues:

1 - To determine if a firearm and toolmark examiner has the ability to correctly associate test fired bullets to the correct consecutively rifled gun barrels;

2 - To expand the test data base from the original 67 participants to participants in laboratories worldwide;

3 - To provide test sets of known bullet pairs and unknown test bullets from the 10 consecutively rifled barrels for laboratories to use in their organizational training programs;

4 - To evaluate the issue of subclass characteristics on bullets fired from consecutively rifled barrels;

5 - To provide information to counter various legal challenges concerning the ability of firearm and toolmark examiners to identify bullets to firearms;

6 - To provide examiners with examples of best known non-match (KNM) bullets.

Materials and Methods

1. Pistol: One Ruger P-85 9mm Luger caliber semiautomatic pistol, serial number: 302-06291 with one 15-cartridge capacity magazine. The same magazine was used during the test firing sequence.

2. Barrels: Ten consecutively rifled 9mm caliber barrels manufactured by Ruger for the Ruger P-85 pistol. The barrels were marked 11 through 20, hereafter referred to as barrel numbers 1 through 10.

3. Ammunition: Winchester 9mm caliber NATO, 124 grain FMJ ammunition, lot number: Q4312, Headstamp: WCC96.

4. Recovery system: One locally manufactured and vented 800 gallon water recovery tank, located in the firearm section of the Indianapolis-Marion County Forensic Services Agency (IMCFSA), Indianapolis, Indiana.

5. Ear and eye protection for test firing, electric engraver unit for scribing test bullets.

6. Envelopes of different sizes, computer labels for labeling the test envelopes, padded packaging materials, pill boxes for collecting test fired bullets, and shipping containers.

Methods (Test Construction)

Each test set included a control set and an unknown set of bullets. In the control set, it was known which barrel fired the bullet and was comprised of two bullets fired from each of the 10 barrels. The unknown set of fifteen bullets was comprised of at least one bullet from each barrel and no more than three bullets from any one barrel. A total of 240 such test sets were prepared.

Prior to test firing the ammunition to prepare the test sets, the pill box containers were appropriately marked to indicate both barrel number and sequence of seven shots. For example, a container marked 1/1 would indicate barrel 1, test sequence 1, while a container marked 7/239 would indicate barrel 7, test sequence 239. Test firing commenced on July 8, 1999 and concluded on August 10, 2000 and was carried out by Hamby, Brundage, and Mickey French, all qualified firearm and toolmark examiners then employed at the IMCFSA. Production of the test ultimately involved shooting some 16,800 cartridges; 1,680 from each of the 10 consecutively manufactured barrels. All 16,800 fired cartridge cases were test fired using the same slide installed on the Ruger P-85 semiautomatic pistol.

Seven cartridges were test fired for each test sequence. The test fired bullets were retrieved from the water recovery tank and placed, along with the recovered cartridge cases, into a pill box designed to maintain them. After the test firing was complete for a group of test sets, the marked pill boxes were combined into 'groups' by barrel and firing sequence number. This process allowed for the same relative amount of barrel wear because the bullets were test fired during the same sequence. For example, every barrel – one through ten – and sequence 74 were assembled into one test set, 1/74, 2/74, 3/74, etc.

The sets of 20 'known' bullets were scribed on the base with the barrel number from 1 to 10. The 15 'unknown' bullets were scribed on the base with an alpha designator from A through Z. To ensure a random letter process and to preclude using the same alpha character twice while scribing the 'unknown' bullets, a set of 3x5 cards were marked A through Z. The 26-card set was shuffled just before scribing the 15 'unknown' bullets and the first 15 alpha characters selected were utilized for marking the bullets. Once the test fired bullets were marked, they were placed into coin envelopes that were previously labeled as depicted below:

KNOWN TWO (2) TEST BULLETS FIRED FROM BARREL #10

QUESTIONED ONE (1) UNKNOWN FIRED BULLET – MARKED 'J'

The test sets were individually packaged according to the sequence of the test set being fired and continued until all 240 test sets were completed. A 10% random sampling of the 240 prepared sets was conducted before the sets were shipped to participants. This random sampling, using an optical comparison microscope, validated that it was possible to identify the 15 'unknowns' to the 'known' bullets.

Each completed test set was sealed in a manila envelope with instructions for completing the examination. The answer sheet requested additional information from the participant, such as years of experience, years and type of training, type of comparison microscope used and membership in forensic organizations. It may be, if the error rate was non-zero, that this could be correlated with training, experience and/or type of microscope. The test materials and answer form were all packaged in a padded envelope for shipment. When the answer form was received from a participant, the answers were evaluated using the test set answer key. A letter of acknowledgement and the answer key were mailed to the participant for later use within their laboratory.

Distribution of Tests

In the expanded study, notices of the tests availability were widely distributed. A letter announcing the availability of the test sets was distributed at the Annual AFTE Training Seminars held in Virginia, Missouri and California in 1999, 2000 and 2001. The test sets were also distributed at the 2000, 2001 and 2002 Shooting, Hunting, Outdoor, Trade (SHOT) Shows. An announcement concerning the availability of the 10 barrel test was also published in the AFTE Journal and the authors contacted a number of individuals – in laboratories in the United States and overseas – to solicit participation in the project. To date, all 240 test sets have been distributed to forensic laboratories, universities and researchers around the world.

Results and Discussion

Test Series	No. Examiners Reporting All Correct Results	No. Inconclusive Results (Examiners, bullets)		No. Incorrect Results
Brundage	66	1	1	0
Hamby	436	4	7	0
Combined Totals	502	5	8	0

A total of 507 responses have been received from individuals that participated in the two studies. In the original Brundage study, one laboratory reported an inconclusive result in that they were unable to associate an unknown bullet with the known bullets due to damage to the projectile. While they reported their finding on one of the 15 bullets as "inconclusive", it would perhaps have been more appropriate to have been reported as "unsuitable". In the expanded study by Hamby & Thorpe, two examiners felt that there were insufficient individual characteristics on two of the bullets due to tank rash [75]. In another instance, two firearm and toolmark examiner trainees were unable to correctly associate 5 of the unknown bullets (1 for one trainee, 4 for the second trainee). In each instance, the participants reported their findings as inconclusive and at no time were misidentifications reported.

In addition to individuals examining the test sets using optical comparison microscopy, five test sets were examined using 'ballistics' imaging equipment. The test sets were examined using the following systems with correct answers reported by the participants. This information indicates that these systems – when properly used – can provide appropriate data:

- Intelligent Automation's SciClops™ - Dr. Ben Bachrach (Maryland, United States);
- Automated Land Identification System (ALIS) - Mr. Tsuneo Uchiyama (Tokyo, Japan);
- Integrated Ballistics Identification System (IBIS) ™ – Mr. Robert Thompson (California, United States);
- BulletTRAX-3D™ - Forensic Technology Scientists (Montreal, Canada) (2 sets)

Evaluation

The majority of participants reported that the examination of the test set required between seven and nine hours. The shortest amount of time reported was three hours while the longest time required for two participants was 30 hours.

In this type of testing, once a bullet is ascribed to a barrel, that bullet is not re-examined; this is sampling without replacement. Normally the probability of achieving a correct result by pure chance is calculated using the hypergeometric theorem. However, this situation is complicated by having up to three separate bullets ascribed to one barrel in a test set and the exact probability will vary depending where in the sequence of fifteen test bullets the additional bullets occur. Therefore, a simpler calculation was used. If an examiner took an "unknown" bullet and attributed it at random to a barrel then there would be a probability of 0.1 that the attribution would be correct. In the survey, each examiner attributed 15 bullets to the 10 barrels correctly and the probability of achieving this by chance is 1 in 10 (-16).

Background information provided from the questionnaires provided insight (shown below) about 435 individuals responding to the survey as the data wasn't available for the original 67 participants or for the individuals using the SciClops™ and BulletTRAX-3D™ imaging systems. Responses were obtained from 20 countries on four continents. Participants from the following countries contributed to this worldwide research project: Australia, Barbados, Belgium, Botswana, Canada, China, Germany, Greece, Jamaica, Japan, Saudi Arabia, Netherlands, New Zealand, Norway, Switzerland, South Africa, Trinidad & Tobago, United Arab Emirates, United Kingdom and the United States. In the United States, responses were received from examiners in 49 states and the territories of Guam and Puerto Rico.

The median number of years of experience in the field, for the 435 respondents participating in the project was 10.5 years, with the amount of time spent in training 1.8 years. Two of the participants were in training and had a less than three months experience each while one individual was a graduate student in a forensic science program. The majority, in excess of 95%, of all responding participants indicate that they were trained under an 'on the job' (OJT) training scheme while a few examiners stated that their training was formal. The larger laboratory systems such as the Federal Bureau of Investigation (FBI) and the Illinois State Police (ISP) – as well as some other laboratories - conduct more formal training than some smaller laboratories. It should be noted, however, that the majority of forensic laboratories around the world utilize a combination of training methods which includes the AFTE Training Manual, specific OJT training, formal instruction, tours of manufacturing facilities, etc. A recent web based firearm and toolmark examiner training program – sponsored by the National Institute of Justice (NIJ) and prepared by the National Forensic Science & Technology Center (NFSTC) under contract to NIJ – was released for use by examiners worldwide at the 2008 AFTE Annual Training Seminar in Honolulu, Hawaii. The majority of the program was written by experienced firearm and toolmark examiners – all members of the Association of Firearms and Toolmark Examiners - and closely follows the AFTE Training Manual [76].

When asked about the Specialized Firearms Techniques School offered by the Federal Bureau of Investigation, a total of 65 participants responded that they had attended the school. Since the National Firearms Examiner Academy (NFEA) was established in 1999, a total of 88 firearm and toolmark examiners have successfully completed the course. Of those attending the NFEA, a total of 21 firearm and toolmark examiners participated in this research project.

Conclusions

A total of 507 responses were received for this worldwide research project, including the 67 responses from firearm and toolmark examiners who participated in the original study by Brundage. The project was designed to determine if trained firearm and toolmark examiners could identify 15 'unknown' fired bullets to the correct one of 10 consecutively rifled barrel. In only two instances of the 7,605 'unknown' fired bullets examined, respondents considered three of the bullets as unsatisfactory for microscopic examination due to damage. Two firearm and toolmark examiner trainees were simply unable to ascribe five of the 'unknown' fired bullets to the 'known' samples. The remaining 7,597 'unknown' fired bullets were correctly identified by participants to the provided 'known' bullets. The fact that there were no actual errors shows that the test procedure used to ascribe bullets fired from consecutively rifled barrels is reproducible on a worldwide basis.

In a Daubert Hearing (a legal challenge in the United States), an examiner could state something like the following: "A long term internationally administered validity test using consecutively rifled barrels, a condition widely considered the most likely to produce errors, was completed by 507 different participants (502 examiners, 5 using instrumentation) and resulted in 7,597 correct identification conclusions and no false positive conclusions".

This study shows that there are identifiable features on the surfaces of bullets that can link them to the barrel that fired them. Although one would expect bullets fired from consecutively rifled barrels to display subclass characteristics, <u>the issue of subclass characteristics was not an issue</u> for the 502 individuals who participated in this research project. Based on the results of this research, having fired bullets in good condition and properly trained firearm and toolmark examiners, the identification process has an extremely low estimated error rate. In circumstances where bullets are deformed or fragmented, the comparison process may be more difficult and the error rate may increase. This study also shows that various statements made about the inability of examiners to associate fired bullets to consecutively rifled barrels were incorrect.

It should be noted that 502 participants – excepting those utilizing 'ballistics' imaging equipment – conducted the examinations using conventional optical comparison microscopy. Results of this study have provided the forensic science community with additional supportive documentation in the field of firearm and toolmark identification, especially as it pertains to the identification of bullets fired from consecutively rifled barrels.

Acknowledgments

The authors gratefully acknowledge the participation of everyone that submitted their data for this research project.

References

1. Brundage, J., "The Identification of Consecutively Rifled Gun Barrels," <u>AFTE Journal</u>, Vol. 30, No.1, Summer 1998, pp. 438 – 444.

2. Hamby, J., Forensic Firearms Examination, Chapter 3, Ph.D. Thesis at the University of Strathclyde, Glasgow, Scotland, UK, 2001.

3. Hamby, J., "The Identification of Consecutively Rifled 9mm Pistol Barrels: A Pre-publication Update" Presented at the 38th AFTE Training Seminar, San Francisco, CA, May, 2007.

4. Hamby, J., "The Identification of Bullets Fired from 10 Consecutively Rifled RUGER 9mm Pistol Barrels –A Research Project Involving 460 Forensic Scientists from Laboratories in 18 Countries" Presented at the 18th Triennial Meeting of the International Association of Forensic Sciences, New Orleans, LA, July, 2008.

5.Personal observation of one of the authors who has had the opportunity to visit over 175 forensic laboratories located around the world during the past 38 years.

6. Moran, B., "A Report on the AFTE Theory of Identification and Range of Conclusions for Toolmark Identification and Resulting Approaches to Casework," <u>AFTE Journal</u>, Vol. 34, No. 2, Spring 2002, pp. 117-235.

7. Morland, N., <u>Science in Crime Detection</u>. Emerson Books Incorporated, New York: 1960.

8. Garrison, D., "Guns of Brownsville," <u>AFTE Journal</u>, Vol. 18, No. 4, October 1986, pp. 65-71.

9. Dougherty, P., 'Report on Two Early United States Firearms Identification Cases", <u>Journal of Forensic Sciences</u>, Vol. 14, No. 4, October 1969, pp. 453-459.

10. Warniment, D., "Brownsville Investigation – A Historical Case Comparison," <u>AFTE Journal</u>, Vol. 34, No. 4, Fall 2003, pp. 375.

11. Hastings, M., The Other Mr. Churchill, Dodd, Mead & Company, London 1965.

12. Lucas, A., Forensic Chemistry and Scientific Criminal Investigation. Edward Arnold & Company, 1931.

13. Goddard, C., "History of Firearms Identification," AFTE Journal, Vol. 21, No. 2, April 1989, pp. 263-278.

14. Bonfanti, M., and De Kinder, J., "The Influence of the Use of Firearms on their Characteristics Marks," AFTE Journal, Vol. 31, No. 1, Summer 1999, pp. 318-323.

15. Nichols, R., "Firearm and Toolmark Identification Criteria: A Review of the Literature", Journal of Forensic Sciences, Vol. 42, No. 3, May 1997, pp. 466-474.

16. Nichols, R., "The History of Firearms and Tool Mark Identification Criteria," Presented at the 55th AAFS Annual Seminar, Chicago, IL February 2003.

17. Nichols, R., "Firearm and Toolmark Identification Criteria: A Review of the Literature, Part II", Journal of Forensic Sciences, Vol. 48, No. 2, March 2003, pp. 318-327.

18. Grzybowski, R., and Murdock, J., "Firearm and Toolmark Identification – Meeting the Daubert Challenge," AFTE Journal, Vol. 30, No. 1, Winter 1998, pp. 3-14.

19. Grzybowski, R., Miller, J., Moran, B., Murdock, J., Nichols, R., and Thompson, R., "Firearm / Toolmark Identification: Passing the Reliability test under Federal and State Evidentiary Standards," AFTE Journal, Vol. 35, No. 2, Spring 2003, pp. 209-241.

20. Biasotti, A., Murdock, J., and Moran, B., Chapter 36, Vol. 4, Modern Scientific Evidence: The Law and Science of Expert Testimony, (Faigman, D., Kaye, D., Saks, M., Sanders, J., and Chen, E., eds., 2006 – 2007), St. Paul: Thompson-West

21. Howe, W., & Munhall, B., "Report on the Formation of the Association of Firearm and Tool Mark Examiners," AFTE Journal, Vol. 31, No. 3, Summer 1999, pp. 219-220 (This article was originally published in the first AFTE publication - AFTE Newsletter, NL #1, May 1969).

22. Hamby, J. and Thorpe, J., "The History of Firearm and Toolmark Identification," AFTE Journal, Vol. 31, No. 3, Summer 1999, pp. 266-284.

23. Berg, S., "The History of Firearms Identification", Identification News, Vol. 15, No. 6, June 1965, pp. 5-15.

24. Berg, S., "The Drama of Forensic Ballistics," AFTE Journal, Vol. 11, No. 3, July 1979, pp. 44-48.

25. Biasotti, A., "The Principles of Evidence Evaluation as Applied to Firearms and Tool Mark Identification," Journal of Forensic Sciences, Vol. 9, No. 4, October 1964.

26. Moran, B., "Reference Collections – The Lifeblood of the Firearm and Toolmark Examiner," AFTE Journal, Vol. 33, No. 2, Summer 2001, pp. 219-226.

27. Hamby, J. and Thorpe, J., "A Historical Perspective of Firearms Reference Collections: Their Size, Composition and Uses," AFTE Journal, Vol. 31, No. 3, Summer 1999, pp. 291-297.

28. Biasotti., A., "Bullet Bearing Surface Composition and Rifling (Bore) Conditions as Variables in the Reproduction of Individual Characteristics on Fired Bullets," AFTE Journal, Vol. 13, No. 2, April 1981, pp. 94-102.

29. Nichols, R., "Firearm and Toolmark Identification: The Scientific Reliability and Validity of the AFTE Theory of Identification discussed within the Framework of a Study of 10 Consecutively Manufactured Extractors," AFTE Journal, Vol. 36, No. 1, Winter 2004, pp. 67-88 and Vol. 36, No. 2, Spring 2004, pp.125.

30. Biasotti, A., "A Statistical Study of the Individual Characteristics of Fired Bullets," Journal of Forensic Sciences, Vol. 4, No. 1, January 1959, pp. 34-50.

31. Thompson, E. and Wyant, R., "Knife Identification Project (KIP)," AFTE Journal, Vol. 35, No. 4, Fall 2003, pp. 366-370.

32. Goddard, C., "The Valentine's Day Massacre: A Study in Ammunition Tracing," AFTE Journal, Vol. 12, No. 1, January 1980, pp. 44-59 (This article originally appeared in the American Journal of Police Science, January-February 1930).

33. Flynn, E., "Toolmark Identification," Journal of Forensic Sciences, Vol. 2, No. 1, January 1957, pp. 95-106.

34. Kirby, S., "Comparison of 900 Consecutively Fired Bullets and Cartridge Cases from a .455 Caliber S & W Revolver," AFTE Journal, Vol. 15, No. 3, July 1983, pp.113-125.

35. Lutz, M., "Consecutive Revolver Barrels," AFTE Newsletter, No. 9, August 1970, pp. 24-28.

36. Hamby, J., "Identification of Projectiles," AFTE Journal, Vol. 6, No.5&6, October / December 1974, p. 22.

37. Murdock, J., The Effects of Crowning on Gun Barrel Individuality. AFTE Newsletter, No. 7, April 1970, pp. 12-13.

38. Personal communication between John Ward and one of the authors.

39. Butcher, S. and Pugh, D., "A Study of Marks made by Bolt Cutters," Journal of the Forensic Science Society, Vol. 15, No. 2, April 1975, pp. 115-126.

40. Ogihara, Y., Kubota, M., Sanada, M., Fukuda, K., Uchiyama, T. and Hamby, J., "Comparison of 5000 Consecutively Fired Bullets and Cartridge Cases From a 45 Caliber M1911A1 Pistol," AFTE Journal, Vol. 15, No. 3, July 1983, pp. 127-140.

41. Watson, D. "The Identification of Toolmarks produced from consecutively manufactured knife blades in soft plastics." AFTE Journal, Vol. 10, No. 3, July 1978, pp. 43-45.

42. Cassidy, F. "Examination of Toolmarks from Sequentially Manufactured Tongue and Groove Pliers," Journal of Forensic Sciences, Vol. 25, No. 4, October 1980, pp.796-809.

43. Murdock, J., "A General Discussion of Gun Barrel Individuality and an Empirical Assessment of the Individuality of Consecutively Rifled 22 Caliber Rifles," AFTE Journal, Vol. 13, No. 3, July 1981, pp. 84–111.

44. Tuira, Y., "Tire Stabbing with Consecutively Manufactured Knives," AFTE Journal, Vol. 14, No. 1, January 1982, pp. 50-52.

45. Hall, E., "Bullet Markings from Consecutively Rifled Shilen DGA Barrels," AFTE Journal, Vol. 15, No. 1, January 1983, pp. 33-53.

46. Shem, R., and Striupaitis, P., "Comparison of 501 Consecutively Fired Bullets and Cartridge Cases from a 25 Caliber Raven Pistol," AFTE Journal, Vol. 15, No. 3, July 1983, pp. 109-112.

47. Matty, W., and Johnson, T., "A Comparison of Manufacturing Marks on Smith & Wesson Firing Pins," AFTE Journal, Vol. 16, No. 3, July 1984, pp. 51-56.

48. Matty, W., "Raven .25 Automatic Pistol Breech Face Tool Marks," AFTE Journal, Vol. 16, No. 3, July 1984, pp. 57-60.

49. Matty, W., "A Comparison of Three Individual Barrels Produced from One Button-Rifled Barrel Blank," AFTE Journal, Vol. 17, No. 3, July 1985, pp. 64-69

50. Van Dijk, T., "Steel Marking Stamps: Their Individuality at the Time of Manufacture," Journal of the Forensic Science Society, Vol. 25, No. 4, July/Aug 1985, pp. 243-253.

51. Uchiyama, T., "Similarity among Breech Face Marks Fired from Guns with Close Serial Numbers," AFTE Journal, Vol. 18, No. 3, July 1986, pp. 15-52.

52. Gross, K., "The 'Hammer-Murderer,'" AFTE Journal, Vol. 27, No. 1, Winter 1995, pp. 27-30.

53. Schecter, B., Silverwater, H. and Etzion, M., "Extended Firing of a GALIL Assault Rifle," AFTE Journal, Vol. 24, No. 1, January 1992, pp. 37-46.

54. Hall, J. "Consecutive cuts made by bolt cutters and their effect on identification," AFTE Journal, Vol. 24, No. 3, July 1992, pp. 260-267.

55. Thompson, E., "Phoenix Arms (Raven) Breechface Toolmarks," AFTE Journal, Vol. 26, No. 2, April 1994, pp. 134-135.

56. Brown, C. and Bryant, W., "Consecutively Rifled Gun Barrels Present in Most Crime Labs," AFTE Journal, Vol. 27, No. 3, July 1995, pp. 254-258.

57. Thompson, E., "False Breechface ID's," AFTE Journal, Vol. 28, No. 2, April 1996, pp. 95-96.

58. Tulleners, F. and Guisto, M., "Striae Reproducibility on Sectional Cuts of One Thompson Contender Barrel," AFTE Journal, Vol. 30, No. 1, Winter 1998, pp. 62-81.

59. Tulleners, F. and Hamiel, J., "Sub Class Characteristics of Sequentially Rifled .38 Special S&W Revolver Barrels," AFTE Journal, Vol. 31, No. 2, Spring 1999, pp. 117-122.

60. Miller, J., "An Examination of Two Consecutively Rifled Barrels and a Review of the Literature," AFTE Journal, Vol. 32, No. 3), Summer 2000, pp.259-270.

61. Rosati, C., "Examination of Four Consecutively Manufactured Bunter Tools," AFTE Journal, Vol. 32, No. 1, Winter 2000, pp. 49-50.

62. Lopez, L. and Grew, S., "Consecutively Machined Ruger Bolt Faces," AFTE Journal, Vol. 32, No. 1, Winter 2000, pp.

19-24.

63. Hamby, J., Forensic Firearms Examination, Chapter 4, Ph.D. Thesis at the University of Strathclyde, Glasgow, Scotland, UK, 2001.

64. Doelling, B., "Comparison of Four Thousand Consecutively Fired, Steel Jacketed Bullets from a 9X18mm Makarov Pistol," Presented at the Annual Meeting of the American Academy of Forensic Sciences, February, 2001.

65. Miller, J., "An Examination of the Application of the Conservative Criteria for Identification of Striated Toolmarks Using Bullets Fired from Ten Consecutively Rifled Barrels," AFTE Journal, Vol. 33, No. 2, Spring, 2001, pp. 125-132.

66. Eckerman, S., "A Study of Consecutively Manufactured Chisels." AFTE Journal, Vol. 34, No. 4. Fall 2002, pp. 379-390.

67. Lee, S., "Examination of Consecutively Manufactured Slotted Screwdrivers," AFTE Journal, Vol. 35, No. 1, Winter 2003, pp. 66-70.

68. Thompson, E. and Wyant, R., "Knife Identification Project (KIP)," AFTE Journal, Vol. 35, No. 4, Fall 2003, pp. 366-370.

69. Bunch, S. and Murphy, D., "A Comprehensive Validity Study for the Forensic Examination of Cartridge Cases," AFTE Journal, Vol. 35, No. 2, Spring 2003, pp. 201-203.

70. Vinci, F., Falamingo, R., Campobasso, C. and Bailey, J., "Morphological Study of Class Characteristics Produced by Firing 2500 Cartridges in a .45 Caliber Semi-Automatic Pistol," AFTE Journal, Vol. 37, No. 4, Fall 2005, pp. 368-372.

71. Clow, C., "Cartilage Stabbing with Consecutively Manufactured Knives: A Response to Ramirez v. State of Florida," AFTE Journal, Vol. 37, No. 2, Spring 2005, pp. 86-116.

72. Smith, E., "Cartridge Case and Bullet Comparison Validation Study with Firearms Submitted in Casework," AFTE Journal, Vol. 37, No. 4, Spring 2005, pp. 130-135.

73. Collins, R., "How "Unique" are Impressed Toolmarks? – An Empirical Study of 20 Worn Hammer Faces," AFTE Journal, Vol. 37, No. 4, Fall 2005, pp. 252-292 (With addendum by Rocky S. Stone)

74. Gouwe, J., Hamby, J., Norris, S., "Comparison of 10,000 Consecutively Fired Cartridge Cases from a Model 22 GLOCK .40 S&W Caliber Semiautomatic Pistol," AFTE Journal, Vol. 40, No. 1, Winter 2008, pp. 57-63.

75. Tank rash – an unofficial term used by some firearms examiners to denote the damage caused when fired bullets strike the bottom of the water recovery tank.

76. www.ojp.usdoj.gov/nij/training/firearms-training/

Validation of Obturation Marks in Consecutively Reamed Chambers

By: Brian Mayland and Caryn Tucker, Illinois State Police, Chicago, Illinois, U.S.A.

Keywords: chamber marks, consecutively manufactured chambers, fireformed marks, obturation marks

ABSTRACT

The identification of fired cartridge cases to a particular firearm based on breech face markings and firing pin impressions has long been accepted in the field of firearms identification. Obturation marks have also been used to make identifications on fired cartridge cases; however, these types of patterns have not been validated as being reproducible and reliable. A nationwide study of obturation marks was conducted using test shots and unknowns from ten consecutively reamed chambers from three different manufacturers. The results show that obturation marks can be used to individualize a firearm.

Introduction

The sidewall of a cartridge case may contain patterns of striations which can be used to determine if a fired cartridge case was fired from a particular firearm. This pattern of markings has been called fireformed chamber striations, or obturation marks. These marks are caused by the expansion of the cartridge case to the chamber walls via the production of gases from the deflagration of the powder in the cartridge and the subsequent extraction of the fired cartridge case after firing. There has not previously been a validation study focused on the chambers of firearms and the obturation marks that they can produce. An article was written in 1987 by Robert Shem regarding obturation marks on rimfire cartridge cases. This article discussed obturation marks as "fireformed chamber striations"1. This article was not a validation study, but rather was an informative article on these types of marks on rimfire cartridge cases.

The purpose of this study was to identify if two or more fired cartridge cases could be identified as being fired from a particular chamber based only on their obturation marks. The chambers used in this study were consecutively reamed.

The following questions were developed for this research:

1. Can trained forensic scientists in the field of firearms identification successfully identify discharged cartridge cases that have been fired from consecutively reamed chambers based only upon obturation marks?

2. Are these obturation marks reproducible and reliable?

Date Received: September 30, 2011
Peer Review Completed: January 17, 2012

Methods and Materials

Three manufacturers were used for this study: Hi-Point Firearms, Kel-Tec Industries, and Sturm, Ruger & Co. Inc. These manufacturers were chosen based on their willingness to participate and their manufacturing processes. A tour of Hi-Point Firearms in Mansfield, Ohio was provided prior to conducting this study. During this tour, the manufacturing process of the chamber was observed and ten consecutively manufactured chambers were obtained. The authors were unable to travel to Kel-Tec and Ruger to observe their manufacturing processes; however, both companies provided ten consecutively manufactured chambers for this study.

The manufacturing of Hi-Point chambers is a three step process. The first step is the clearing of the chamber, where metal is removed from the interior of the chamber. This reamer is used approximately 5,000 - 6,000 times. The second step is the roughing of the chamber which shapes the metal to the correct size. This reamer can be used 300 - 400 times. The third step is the finishing step, which is the final polishing of the interior of the chamber. This finishing reamer can be used 400 - 500 times. After the polishing reamer has dulled, it is generally used as the roughing reamer. Each of the reamers can be resharpened or discarded; however, the production cost is very low so they are normally discarded.

The manufacturing of Kel-Tec chambers is a two or three step process. The first step is to use a lathe with a cutting reamer. This reamer can be used approximately 90 - 100 times. The second step is a heat treatment. Upon visual inspection, if the interior of the chamber does not look completely finished they will use a Cratex polishing wheel as the final step. When the authors originally purchased the barrels for this project, the third step was done on every chamber.

The manufacturing of Ruger chambers is a three step process. The first step is the roughing of the chamber. This reamer can be used approximately 200 - 400 times. The second step is a finishing reamer, which can also be used approximately 200 - 400 times. The final step is a roller burnishing process. This method of roller burnishing is a cold rolling process without the removal of metal.

Each chamber was utilized to obtain the test shots and the questioned samples. One firearm from each manufacturer (Hi-Point Model C, Ruger Model P89, and Kel-Tec Model P-11) was used for this study; the only parts replaced were the chambers. The cartridge cases were then microscopically compared to identify them to the proper chamber based on their obturation marks. The breech face, firing pin, extractor, and ejector marks were the same on each sample for each respective firearm and could not be used for identification purposes.

A caveat with regard to utilizing obturation marks is being able to exclude manufacturing marks on the ammunition. Another caveat is being able to exclude chambering marks on the ammunition from cycling the cartridge through the firearm and not from discharging the firearm. Therefore, to identify these types of marks as obturation marks and not toolmarks of an unknown origin, it is beneficial to have the firearm submitted for comparison. This allows the examiner to view the manufacturing marks on the cartridge prior to chambering or test firing the cartridge, to view the cartridge once it has been chambered in the firearm, and to view the cartridge case once it has been discharged from the firearm.

Three different brands of ammunition were used in this study: Winchester, Remington and Federal 9 mm Luger caliber with 115 grain full metal jacket bullets. These manufacturers were chosen because they are very common in the United States and are frequently seen in casework.

Each chamber was fired a total of thirty times (ten Winchester, ten Remington, and ten Federal). In addition, each chamber had nine unfired cartridges cycled through it without firing (three of each brand of ammunition). A total of fifty-four validation tests were made. Each validation test consisted of three unknowns and five sets of standards. The five sets of standards each consisted of two fired cartridge cases, one unfired cartridge that was cycled through the chamber, and one unfired cartridge taken directly from the box of ammunition. This allowed the participant to determine which patterns were caused by the discharge of the firearm, which ones were caused by cycling the cartridge through the chamber of the firearm, and which were caused by the ammunition manufacturing process. All the standards and questioned samples were electrically scribed on the inside of the mouth of the cartridge case. The standards were labeled one through ten for the Ruger firearm, eleven through twenty for the Kel-Tec firearm, and twenty-one through thirty for the Hi-Point firearm, for each chamber in which they were fired. The questioned samples were randomly numbered utilizing a computer-based program.

Results

There were sixty-four participants from nineteen laboratory systems nationwide. Twenty-three Ruger, twenty-four Kel-Tec, and seventeen Hi-Point tests were completed. Fifty-five of the sixty-four participants correctly identified all three unknown samples. Six participants correctly identified two of the three unknown samples and had an inconclusive result on the remaining sample. One participant correctly identified one of the unknown samples and had an inconclusive result on the two remaining samples. One participant had inconclusive results on all three unknown samples. One participant incorrectly identified all three unknown samples. As it pertains to this research, the sensitivity is a calculation of the ratio between correct identifications to actual identifications. A ratio of one would mean that all of the identifications within the series of tests were correctly reported as identifications. In this study, the sensitivity is 178/192, or 0.927. A further breakdown of the results follows in Tables 1 and 2.

Conclusion

While the authors acknowledge that one participant incorrectly identified all three of their unknown samples, the authors also have verified that these three same unknown samples were correctly associated by two other examiners during this research project. The authors have therefore determined that this result is due to participant error. This research verifies that obturation marks are reproducible and reliable and that trained firearms examiners can correctly identify fired cartridge cases based on these obturation marks. This research also verifies that the patterns produced by the chambers of these three manufacturers are distinguishable and unique.

Reference

[1] Shem, R.J., "Fireformed Chamber Striations on Rimfire Cartridge Cases," AFTE Journal, Volume 19, No. 3, July 1987, pp. 282-283.

	Ruger	Kel-Tec	Hi-Point
Total number of tests returned	23	24	17
Number of tests with all three samples correctly identified	21	19	15
Number of tests with two of three samples correctly identified and one inconclusive	1	4	1
Number of tests with one of three samples correctly identified and two inconclusive	0	1	0
Number of tests with three inconclusive	0	0	1
Number of tests with incorrect answers	1	0	0

Table 1

Test #	Number of Times Test Used	Results
Ruger		
1	2	3 ID
		2 ID/1 INC
2	0	
3	1	3 ID
4	2	3 ID
		3 ID
5	1	3 ID
6	1	3 ID
7	3	3 ID
		3 ID
		3 WRONG
8	1	3 ID
9	0	
10	1	3 ID
11	0	
12	2	3 ID
		3 ID
13	2	3 ID
		3 ID
14	1	3 ID
15	2	3 ID
		3 ID
16	1	3 ID
17	1	3 ID
18	2	3 ID
		3 ID

Test #	Number of Times Test Used	Results
Kel-Tec		
19	2	3 ID
		3 ID
20	2	3 ID
		3 ID
21	1	3 ID
22	1	3 ID
23	2	3 ID
		2 ID/1 INC
24	2	2 ID/1 INC
		1 ID/2 INC
25	1	2 ID/1 INC
26	2	3 ID
		3 ID
27	2	3 ID
		3 ID
28	2	3 ID
		2 ID/1 INC
29	1	3 ID
30	0	
31	0	
32	1	3 ID
33	1	3 ID
34	1	3 ID
35	1	3 ID
36	2	3 ID
		3 ID

Test #	Number of Times Test Used	Results
Hi-Point		
37	2	3 ID
		3 ID
38	1	3 ID
39	2	3 ID
		3 ID
40	1	3 ID
41	1	3 INC
42	2	3 ID
		2 ID/1 INC
43	1	3 ID
44	1	3 ID
45	0	
46	0	
47	0	
48	0	
49	2	3 ID
		3 ID
50	0	
51	1	3 ID
52	0	
53	2	3 ID
		3 ID
54	1	3 ID

Table 2

An Empirical Study to Improve the Scientific Foundation of Forensic Firearm and Tool Mark Identification Utilizing 10 Consecutively Manufactured Slides

By: Thomas G. Fadul Jr., Ph.D., Gabriel A. Hernandez, M.S., Stephanie Stoiloff, M.S., Miami-Dade Police Department Crime Laboratory, Miami, Florida; Sneh Gulati, Ph.D, Department of Statistics, Florida International University, Miami, Florida.

Keywords: breechface, cartridge cases, consecutive manufacture, error rate, NAS, Ruger slides, validity

ABSTRACT

The Miami-Dade Police Department Crime Laboratory conducted a research study on The Repeatability and Uniqueness of Striations/Impressions on Fired Cartridge Casings Fired in 10 Consecutively Manufactured 9mm Ruger Slides to improve understanding of the accuracy, reliability, and measurement validity in the firearm and tool mark discipline of forensic science. The foundation of firearm and tool mark identification is that each firearm/tool produces a signature of identification (striation/impression) that is unique to that firearm/tool, and through the examination of the individual striations/impressions, the signature can be positively identified to the firearm/tool that produced it. The National Academy of Sciences (NAS) Report questioned the repeatability and uniqueness of striations/impressions left on fired evidence as well as the validity and error rate in firearms identification. This study analyzed the repeatability and uniqueness of striations/impressions on fired cartridge cases fired in 10 consecutively manufactured Ruger slides by analyzing breech face striations/impressions through an evaluation of the participants' accuracy in making correct identifications. One semi-automatic pistol and nine additional consecutively manufactured slides were utilized. Consecutively manufactured slides are significant to this study because they were manufactured with the same equipment/tools. Even though these slides were consecutively made, their signatures should be different. Test sets assembled included test fired casings from each slide, as well as unknowns. Participants were firearm & tool mark examiners throughout the United States. One hundred and fifty-eight test sets were distributed to laboratories in forty-six states and the District of Columbia. The test sets were designed to determine an examiner's ability to correctly identify cartridge casings fired from 10 consecutively manufactured Ruger Slides to test fired cartridge casings fired from the same slides. This empirical study established an error rate of less than 0.1 percent. Durability testing established that the Ruger Slides maintained their individual signature after multiple firings.

This project was supported by Award No. 2009-DN-BX-K230 awarded by the National Institute of Justice, Office of Justice Programs, U. S. Department of Justice. The opinions, findings, and conclusions or recommendations expressed in this publication/program/exhibition are those of the authors and do not necessarily reflect those of the Department of Justice.

EXECUTIVE SUMMARY

The National Academy of Sciences Report (2009) questioned the repeatability and uniqueness of striations/impressions left on fired firearms evidence as well as the validity and error rate in firearms identification. The goals of this research study were: 1) to conduct an empirical study to evaluate the repeatability and uniqueness of striations/impressions imparted by consecutively manufactured slides; and 2) to determine the error rate for the identification of same gun evidence.

Utilizing an experimental research design, this study analyzed the repeatability and uniqueness of striations/impressions on cartridge cases fired in 10 consecutively manufactured slides by analyzing breech face striations/impressions. One semi-automatic pistol and nine additional consecutively manufactured slides were utilized. Consecutively manufactured slides are significant to this study because they were manufactured with the same equipment/tools. Even though these slides were consecutively made, their signatures should be different if there is no subclass influence. Test sets assembled included known test fired casings from each slide, as well as unknowns.

Participants were firearm & tool mark examiners throughout the United States. One hundred and fifty-eight test sets were distributed to laboratories in forty-six states and the District of Columbia. The test sets were designed to determine an examiner's ability to correctly identify cartridge casings fired from 10 consecutively manufactured Ruger slides to test fired cartridge casings fired from the same slides. This empirical study established an error rate of less than 0.1 percent. Durability testing established that the breech faces

Date Received: March 6, 2013
Peer Review Completed: August 9, 2013

of the Ruger slide maintained their individual signature after multiple firings.

The National Academy of Sciences Report (2009) states that "some forensic science disciplines are supported by little rigorous systematic research to validate the discipline's basic premises and techniques." In addition, the report states that forensic science will be improved by collaborative opportunities "with the broader science and engineering communities." The statistical analyses of the research data was performed by a professor from the Department of Statistics at Florida International University. This collaboration with an external agency to analyze the data that was collected helps to ensure that the statistical results are reported accurately and without bias.

This research provides further support for the scientific foundation of forensic firearm and tool mark identification through the evaluation of unknown cartridge casings to determine the repeatability and uniqueness of striations/ impressions. The error rate of identifications of same gun evidence was calculated from the data collected. This study provides empirical data to strengthen the foundation of firearms identification in both the firearm identification field and in the legal arena, thus addressing some of the National Academy of Sciences' concerns with how firearm and tool mark identifications are supported.

INTRODUCTION

Purpose, Goals and Objectives

The purpose of this research was to conduct an empirical study to evaluate the repeatability and uniqueness of striations/ impressions imparted by consecutively manufactured slides to fired cartridge casings as well as to determine the error rate for the identification of same gun evidence.

Limited studies have previously been conducted on consecutively manufactured slides; however, no studies have been conducted in which test sets were sent to a large number of participants. The goal was to determine whether or not fired cartridge casings can be identified to the firearms that fired them through the comparison of tool marks.

The objective of this research study was to determine if firearm and tool mark examiners would be able to identify unknown casings to the firearms that fired them when examining casings fired through consecutively manufactured slides utilizing individual, unique and repeatable striations/ impressions. Also, the study presented herein evaluated the experience level of firearm and tool mark examiners and the effect of their experience level on the results.

Review of Relevant Literature

A review of the relevant literature provides a multitude of examples of studies where one individual or a small group of individuals correctly identified casings fired consecutively in various firearms. These durability studies examined the repeatability and longevity of a firearm's unique signature. These studies are conducted by comparing the first test fire to the last test fire after the firearm has been fired at least several times in between. The research has demonstrated that a firearm's unique signature remains identifiable even after several thousand test fires.

Ogihara et al. (1983) examined 5,000 fired casings fired through one .45 Auto caliber pistol and correctly identified all of the fired casings to this pistol. Shem and Striupaitis (1983) conducted a durability study with one .25 Auto caliber pistol and reported that they were able to correctly identify the first fired casing to the 501st fired casing. Matty (1984) examined casings that were fired from three consecutively manufactured .25 Auto caliber slides and concluded that they could be identified with the correct slide. Thompson (1994) reported that he examined casings fired from four consecutively manufactured .25 Auto caliber slides and concluded that they could also be identified to the correct slides.

Hamby (2001) examined and correctly identified casings to the 617 firearms that fired them. Bunch and Murphy (2003) conducted a study at the FBI Laboratory utilizing 10 consecutively manufactured slides; they concluded that each slide produced a different signature. Coody (2003) examined and correctly identified fired casings from 10 consecutively manufactured Ruger pistol slides. Coffman (2003) examined and correctly identified fired casings from 10 consecutively manufactured breech bolts. Vinci et al. (2005) conducted a durability study utilizing one pistol and determined that they could correctly identify all 2,500 fired casings. Gouwe et al. (2008) conducted a durability study with one .40 caliber firearm and reported that they were able to identify fired casing one to fired casing 10,000.

No study has been conducted to identify casings from consecutively manufactured slides where the number of participants in the study reaches or exceeds one hundred. A larger sample size will lead to a more reliable estimate of the true error rate for the identification of same gun evidence by firearm and tool mark examiners.

RESEARCH DESIGN AND METHODS

This study utilized an experimental research design (Christensen, 2004; Creswell, 2005), and was conducted in a crime laboratory setting. Participants compared questioned casings to known standards that were fired in 10 consecutively manufactured slides in order to determine whether or not the consecutively manufactured slides differed from each other by producing different signatures, each with unique striations/impressions (tool marks). Durability testing was then conducted to determine if the individual signature of the slides changed due to repeated firing. This research study also established an error rate for the identification of same gun evidence.

Quantitative data was utilized to determine if the examiners could correctly identify questioned casings test fired in multiple consecutively manufactured slides. Additionally, the years of experience of the examiners was recorded. This data answered the following: 1) whether or not consecutively manufactured slides produced different individual signatures; 2) whether multiple firings changes the signature to the extent where it can no longer be identified; and 3) whether years of experience impacts correct identifications. Questionnaire/answer sheets were utilized to collect the quantitative data (see Appendices A and B).

The proposed outcome in this section is presented with the intention that the findings will be able to answer the research questions.

Research Questions

Q1. Will firearm and tool mark examiners be able to correctly identify the firearms that fired the questioned casings when examining casings fired through consecutively manufactured slides?

Q2. Will firearm and tool mark examiners with less than 10 years of experience reach the same conclusions as those with greater than 10 years of experience when examining casings fired through consecutively manufactured slides?

Q3. Will firearm and tool mark examiners be able to correctly identify the firearms that fired the questioned casings when examining casings fired at different intervals (durability study - after 361 to 995 test firings – see Durability Phase 2 Testing Study Definition on page 383) through consecutively manufactured slides?

Research Hypotheses

H1. Firearm and tool mark examiners will be able to correctly identify unknown casings to the firearms that fired them when examining casings fired through consecutively manufactured slides by utilizing individual, unique and repeatable striations/impressions.

H2. The experience level of firearm and tool mark examiners will not affect identification results when examining casings fired through consecutively manufactured slides.

H3. Firearm and tool mark examiners will be able to correctly identify unknown casings to the firearms that fired them when examining casings fired at different intervals through consecutively manufactured slides by utilizing individual, unique and repeatable striations/impressions.

There is one dependent variable that was examined in this study. The dependent variable is accuracy (proportion of incorrect identifications), which was measured by whether or not the questioned casings could be correctly identified to the consecutively manufactured slides by using individual, unique and repeatable striations/impressions.

There are several independent variables in this study, such as the consecutively manufactured slides, interval of firing and experience of the examiner. For Q1 and H1, the researchers were interested in studying the effect of the consecutively manufactured slides on the ability to identify same gun evidence. For Q2 and H2, the researchers were interested in studying the effect of the independent variable of experience, the knowledge and practical wisdom gained through study, observation, experimentation and case work, on the ability to identify same gun evidence. For Q3 and H3, the researchers were interested in studying the effect of the interval of firing on the ability to identify same gun evidence. Extraneous variables were controlled as much as possible by utilizing laboratory settings.

Using an experimental design, three research questions were explored in this study. Question one: Will firearm and tool mark examiners be able to correctly identify the firearms that fired the questioned casings when examining casings fired through consecutively manufactured slides? For question one, the dependent and independent variables were measured through the average error rate on the Consecutively Manufactured Slide Test Set Instrument Survey by a 1 to 15 point system (1 point for each correct answer, with a maximum point value of 15).

Question two: Will firearm and tool mark examiners with less than 10 years of experience reach the same conclusions as those with greater than 10 years of experience when examining casings fired through consecutively manufactured slides? For question two, the dependent and independent variables were measured through the average error rate on the Consecutively Manufactured Slide Test Set Instrument Survey by a 1 to 15 point system (1 point for each correct answer, with a maximum point value of 15).

Question three: Will firearm and tool mark examiners be able to correctly identify the firearms that fired the questioned casings when examining casings fired at different intervals (durability study) through consecutively manufactured slides? For question three, the dependent and independent variables were measured through the average error rate on the Consecutively Manufactured Slide Test Set Instrument Survey Phase 2 by a 1 to 5 point system (1 point for each correct answer, with a maximum point value of 5).

Three hypotheses were tested in this study. For the first hypothesis (H1), the dependent and independent variables measured whether or not consecutively manufactured slides produced individual, unique and repeatable striations/impressions based on each participant's results. If the breech face tool mark signatures from each of the ten consecutively manufactured slides could be distinguished from one another by the participants, this would establish that there is no subclass tool mark influence present from the manufacturing process used to form the breech faces.

For the second hypothesis (H2), the dependent and independent variables measured whether or not the years of experience of the participants affected the ability of the examiners to identify same gun evidence based on each participant's results.

For the third hypothesis (H3), the dependent and independent variables measured whether or not consecutively manufactured slides produced individual, unique and repeatable striations/impressions on casings fired at different intervals based on each participant's results.

Target Population

In this study, the target population represented a subset of the forensic science community, more specifically, firearm and tool mark examiners employed by a law enforcement agency (crime laboratory), or like agency, in the United States. The Miami-Dade Police Department (MDPD) Crime Laboratory (CL) utilized the membership list for the Association of Firearm and Tool Mark Examiners (AFTE). Eleven members of this association currently work in the MDPD CL.

Membership in AFTE is limited to individuals with suitable education, training, and experience in the examination of firearms and/or tool marks. For purposes of this membership, a practicing firearm and/or tool mark examiner is defined by AFTE (2009) as: "(1) An individual who derives a substantial portion of his livelihood from the examination, identification, and evaluation of firearms and related materials and/or tool marks; and (2) An individual whose present livelihood is a direct result of the knowledge and experience gained from the examination, identification, and evaluation of firearms and related materials and/or tool marks."

Every firearms examiner in the United States who is a member of AFTE had an equal opportunity to be included in this study. Each AFTE member was contacted by the MDPD CL via email inviting them to participate in this study, which included completing demographic questions and participation in an experimental exercise. The number of participants exceeded the recommended number based on the formula of $n > 50 + 8m$ (Green, 1991).

The test sets utilized in this study were similar to the work that the participants perform on a routine daily basis. The researchers at the MDPD CL are members of AFTE, and one of the privileges of membership is access to the membership list.

Eligibility-Inclusion Criteria

Participants were required to be firearm and tool mark examiners employed by a law enforcement agency (crime laboratory), or like agency, in the United States, and must have completed a two year training program. Independent examiners who retired from a qualifying agency were also eligible to participate in this study. Participants for the durability testing (questioned casings fired at different intervals) were required to have completed Phase 1 testing of this study.

Accessible Population

Accessibility was limited to firearm and tool mark examiners for whom the MDPD CL was able to obtain email addresses by querying the membership list for AFTE. This accounts for 92% of the 2010 AFTE Roster.

Sampling Plan and Setting

The sampling plan for this study utilized an abstract population. Every eligible firearm and tool mark examiner in the United States who is a member of AFTE, with a functional email address, was invited to participate in this research. The AFTE list identified participants that met the MDPD CL eligibility-inclusion criteria as stated above in Eligibility-Inclusion Criteria. The accessible population included approximately 800 firearm and tool mark examiners in the United States.

To ensure confidentiality, the researchers at the MDPD CL invited firearm and tool mark examiners to participate via email. The survey and test (see Appendices A and B) were conducted by each participant independently, which strengthened the study's validity (Gall & Borg, 1996).

Instrumentation

This study utilized two similar instruments, each of which included two methods of instrumentation: a questionnaire that included the participant's demographics, as well as an answer sheet for an experimental exercise. The questionnaires each took less than ten minutes to complete. The first experimental exercise took approximately two to eight hours to complete. The durability exercise took approximately two to four hours to complete. The above listed approximate times were based on personal communication and observation of participants from the MDPD CL.

The experimental exercise was originally utilized by Brundage (1998), and redesigned by Hamby (2001). Over 500 firearm and tool mark examiners have used this instrument (Hamby, Brundage and Thorpe, 2009). The researchers at the MDPD CL modified this instrument by replacing "barrel" with "slide." In addition, the number of years of experience for each firearm and tool mark examiner was added. Furthermore, the researchers at the MDPD CL added a category for pattern matching and quantitative consecutive matching striations (QCMS). Pattern matching and QCMS are two forms of striated tool mark examination/assessment processes used in the field of firearm and tool mark examination.

The instrument for the durability testing was modified further to include questions about certification and gender. Additionally, the category of "Inconclusive" was added for the experimental exercise.

Data Collection Methods

The researchers did the following:

1. Received National Institute of Justice (NIJ) approval.
2. Sent email to the AFTE membership. Participation was voluntary.
3. Obtained 1 pistol frame and 10 slides and labeled the slides 1 through 10.
4. Obtained 9mm cartridges (ammunition/bullets).
5. Utilized the MDPD CL indoor range for test firing and retrieval of the casings.
6. Placed each slide one at a time on the pistol frame.
7. Loaded the pistol with five cartridges.
8. Fired the pistol on the range.
9. Fired five cartridges through each slide to create one test set. (This was repeated 200 times per slide, 1,000 cartridges per slide in total).
10. Used secure, properly labeled, containers to keep each group of five casings separated.
11. Labeled two of the five casings with the number of the slide in which they were fired (1 through 10) to create the test fired casings (known standards). These known standards were placed in a labeled coin envelope.
12. Labeled remaining three casings with an alpha character designated by the researchers at the MDPD CL to represent the questioned casings (different alpha characters were assigned to each slide).
13. Selected one questioned casing from each slide randomly from the container and placed it in a labeled coin envelope.
14. Selected an additional five questioned casings to complete the test set of 15 questioned casings. These five casings were each placed in a labeled coin envelope.
15. Created 200 test sets and placed each test set in a medium manila envelope.
16. The researchers microscopically examined every 10th set to ensure that the casings were comparable and identifiable.
17. Mailed test sets to respondents. Each respondent received one test packet through the mail which included the following:

    o One questionnaire/answer sheet

    o 15 questioned casings

- o 10 sets of test fired casings (known standards) that were fired through the 10 consecutively manufactured slides.

18. Instructed the participants via the questionnaire/answer sheet to compare the questioned casings to the known standards, and to place their answers on the questionnaire/answer sheet.

    - o The participants were also asked to complete the questions that were on the questionnaire/answer sheet.
    - o The instructions directed the participants to mail the questionnaire/answer sheet or to fax it.

19. Conducted the data collection process for 26 weeks.

20. Utilized an Excel spreadsheet to record and analyze the data collected.

21. Submitted the data to a professor from the Department of Statistics at Florida International University for statistical analyses.

Durability Testing:

22. Sent email to participants who completed the first test. Participation was voluntary.

23. Used 1 pistol frame and 5 slides and labeled the slides.

24. Obtained 9mm cartridges (ammunition/bullets).

25. Utilized the MDPD CL indoor range for test firing and retrieval of the casings.

26. Placed each slide one at a time on the pistol frame.

27. Loaded the pistol with one cartridge.

28. Fired the pistol on the range.

29. Fired one cartridge through each slide to create one test set. (This was repeated 100 times per slide, 100 cartridges per slide in total).

30. Used secured, properly labeled containers to keep each casing separated.

31. Labeled each casing with an alpha character designated by the researchers at the MDPD CL to represent the questioned casing (different alpha characters from the first test were assigned to each slide).

32. Placed each questioned casing in a labeled coin envelope.

33. Created 100 test sets and placed each test set in a medium manila envelope.

34. Mailed test sets to participants. Each respondent received one test packet through the mail which included the following:

    - o One questionnaire/answer sheet
    - o 5 questioned casings
    - o Note: Each participant already had the test fired known standards from the first test.

35. Instructed the participants via the questionnaire/answer sheet to compare the questioned casings to the known standards (already in their possession) and to place their answers on the questionnaire/answer sheet.

    - o The participants were also asked to complete the questions that were on the questionnaire/answer sheet.
    - o The instructions directed the participants to mail the questionnaire/answer sheet or to fax it.

36. Conducted the data collection process for 6 weeks.

37. Utilized an Excel spreadsheet to record the data collected.

38. Submitted the data to a professor from the Department of Statistics at Florida International University for statistical analyses.

Data Coding

Phase 1 Testing

Each participant was assigned a number from 1 to end. There were 15 variables (questioned casings) which were designated with an alpha character and coded as correct (1), incorrect (2) or inconclusive (3). The overall correct number was coded 1 through 15 based on the correct number of identifications. Pattern matching was coded as 1, QCMS was coded as 2, utilization of both methods was coded as 3 and no answer was coded as 4. The type of microscope was coded 1 for Leica, 2 for Leeds, 3 for other, or 4 for no answer. Type of lighting was coded 1 for fluorescent, 2 for fiber optic, 3 for LED, 4 for other, or 5 for no answer. Years of experience was coded based on <10 (coded 1) and >10 (coded 2) years of experience. Examination of other evidence was coded 1 for "yes," 2 for "no" and 3 for no answer. Professional or forensic organizations were coded 1 for "yes," 2 for "no" and 3 for no answer. FBI Specialized Techniques School was coded 1 for "yes," 2 for "no" and 3 for no answer. The number of years of training was coded 1 for 2 years or more and 2 for < 2 years. The type of training was coded into 4 groups, 1 for in-house/structured, 2 for National Firearms Examiner Academy, 3 for other, and 4 for no answer. Individuals trained in QCMS were coded 1 for "yes," 2 for "no" and 3 for no answer (see Appendix A).

Phase 2 Testing (Durability)

Each participant was assigned a number from 1 to end. There were 5 variables (questioned casings) which were designated with an alpha character and coded 1 as correct, 2 as incorrect or 3 as inconclusive. The overall correct number was coded 1 through 5 based on the correct number of identifications. Pattern matching was coded as 1, QCMS was coded as 2, utilization of both methods was coded as 3 and no answer was coded as 4. Type of microscope was coded 1 for Leica, 2 for Leeds, 3 for other, or 4 for no answer. Type of lighting was coded 1 for florescent, 2 for fiber optic, 3 for LED, 4 for other, or 5 for no answer. Gender was coded 1 for male, 2 for female and 3 for no answer. AFTE certification was coded 1 for "yes," 2 for "no" and 3 for no answer. ABC certification was coded 1 for "yes," 2 for "no" and 3 for no answer. Other certification was coded 1 for "yes," 2 for "no" and 3 for no answer (see Appendix B).

Descriptive Analysis

Descriptive analysis was used to describe the participants. Descriptive analysis for Phase 1 and Phase 2 testing included years of experience, method used (pattern matching/QCMS), as well as the type of microscope and lighting used.

Data Analysis Methods

Simple descriptive scores were used to analyze all variables. Next, correlation statistics were performed utilizing a statistical program, S-PLUS, to answer the three research questions. An independent statistician performed the data analyses.

Error Rate Definition

An error rate is a calculated value that represents the comparison of the number of wrong responses with the total number of responses. The error rate for each participant was defined as the proportion of questions answered incorrectly by that participant. For example, if a participant answered 5 out of the 15 questions incorrectly, their error rate is 0.3333. An average error rate is calculated by dividing the sum of the error rates per respondent by the total number of respondents. An average error rate calculation was used for both phase 1 and phase 2 of this study. An average error rate calculation was used by the researchers because it is illustrative of the error rate across all participants rather than solely based on number of responses.

Durability Phase 2 Testing Study Definition

The purpose of a durability test is to evaluate the robustness of repeatable, unique and identifiable striations. Each participant received 5 additional questioned casings with a new answer sheet. The participants were asked to compare these five questioned casings to the known standards that they previously received with the original test set. Each slide had already been fired 1,000 times prior to the 5 additional questioned casings created for the durability study. For example, to create test 1, each slide was used to fire 5 cartridge casings. A total of 995 additional casings were fired through these slides to create the test sets. Therefore, when the durability test set was created for the first test of Phase 2, there were a total of 995 rounds fired through each slide between the creation of the known standards for test 1 and the 5 additional questioned casings fired for the durability study (Phase 2). Each durability test set followed the same sequence.

Internal Validity Strengths

- The quantitative data was internally valid due to the procedures set forth to assemble the tests.
- All the test materials were assembled in a crime laboratory setting.
- All questioned casings and known standard casings were labeled with a number (standard) or letter (questioned casings).
- Secure containers were used to keep the questioned casings separated into groups.
- The researchers at the MDPD CL microscopically examined every 10th test set to ensure that the casings were comparable and identifiable.
- The questionnaire/answer sheet used has been documented in previous studies, and is a standardized format.

Internal Validity Weaknesses

- The validity of this study was dependent upon the accuracy of the assembly of the tests.
- Communication between participants could have threatened the internal validity.
- The possibility exists that the questioned casings and known standards failed to mark clearly. Since every set was not microscopically examined to ensure that the casings were comparable and identifiable, some sets may have contained casings that were not suitable for identification.

External Validity Strengths

- The external validity strength of this research project was that all testing was conducted in a crime laboratory setting.
- Participants utilized a comparison microscope.
- The participants were trained firearm and tool mark examiners.
- The training and experience of the participants strengthened the external validity.
- The researchers exceeded the sample size.

External Validity Weaknesses

- The researchers assumed that the participants followed appropriate AFTE procedures, as listed in the AFTE Procedures Manual, FA-IV-13, Microscopic Comparison (2001).
- The researchers had no control over the equipment used by the participants.
- The training and skill level as well as the experience of the participants could have been an external weakness.
- The participants could have used the well defined firing pin aperture shear striated tool marks since these were adjacent to the breech face marks.
- Circled responses on the Phase 1 answer sheet were marked as correct or incorrect. The Phase 1 answer sheet did not have a designated area to list inconclusive results. However, some examiners did list inconclusive results in the margins of the Phase 1 answer sheet. For Phase 2, a designated area for inconclusive results was present. Correct, not correct, and inconclusive were the three tabulated responses. If no alpha character was selected, it was considered an inconclusive answer. No eliminations were reported on either Phase 1 or Phase 2.
- The participants were not told whether the questioned casings constituted an open or closed set. However, from the questionnaire/answer sheet, participants could have assumed it was a closed set and that every questioned casing should be associated with one of the ten slides.

RESULTS

In this section, the examination of research questions, hypotheses testing, and other findings related to this study were analyzed to evaluate the repeatability and uniqueness of striations/impressions imparted to consecutively manufactured slides as well as to determine the error rate for the identification of same gun evidence. Participant performance (experimental exercise) relating to accuracy and demographic characteristics relating to the participants' ability to perform the experimental exercises were examined.

For this research study regarding participant performance relating to accuracy and methods utilized, a mass email was sent out to the membership of the Association of Firearm and Tool Mark Examiners. A total of 281 examiners representing 157 crime laboratories in 46 states, including the District of Columbia, completed the *Consecutively Manufactured Slide Test Set* questionnaire/answer sheet. Sixty-four of the 281 participants did not meet the two year training requirement for this study. This resulted in a data-producing sample of 217 participants for Phase 1 testing. Additionally, 114 participants completed the Phase 2 testing (Durability) *Consecutively Manufactured Slide Test Set* questionnaire/answer sheet.

The firearm and tool mark examiners that responded to the *Consecutively Manufactured Slide Test Set* questionnaire/answer sheet represented 92% of the states in the United States that conduct firearm and tool mark examinations.

The questionnaire/answer sheet instrument utilized for this study allowed the participants to record their answer by circling the appropriate alpha designator of the unknown casings on the same line as the known test fired casing sets designated by a numerical number 1 – 10 (Brundage, 1998; Hamby, 2001; Hamby & Brundage, 2007, 2009; Fadul 2011). This experimental exercise of the instrument was designed to measure accuracy.

The statistician utilized the statistical analysis program S-PLUS for this study. Nonparametric tests, namely the

Wilcoxon Signed Rank test, the Wilcoxon rank Sum Test and the Kruskall Wallis tests were used for the analysis. The Wilcoxon Signed Rank test is a nonparametric alternative to the paired Student's t-test, the Wilcoxon Rank Sum test is used for comparing two independent samples while the Kruskall Wallis test is used for more than two independent samples. The tests are used when sample populations cannot be assumed to follow a normal distribution. Quite often these tests are based on ranks. As an example, when comparing two independent samples from populations A and B, one would first combine the two samples and rank their tested values (number of incorrect responses, for example) from the lowest to the highest. The lowest observation gets rank 1, the next one rank 2, etc. After ranking the combined sample, one would separate the samples and sum up the ranks of each (say A or B). If the populations are roughly similar, there should be no significant difference in the sum of the ranks (adjusted for sample size). This is the basis for the Wilcoxon Ranked Sum test. The Kruskall Wallis test extends this concept to more than two populations.

The National Academy of Sciences Report (2009) states that "some forensic science disciplines are supported by little rigorous systematic research to validate the discipline's basic premises and techniques." In addition, the report states that forensic sciences will be improved by collaborative opportunities "with the broader science and engineering communities." The statistical analyses of this research data was performed by Dr. Sneh Gulati, a professor from the Department of Statistics at Florida International University. This collaboration with an external agency to analyze the data that is collected helps to ensure that the statistical results are reported accurately and without bias.

Instrument Parameters

Each participant received a total of 10 pairs of known test fired casings labeled Slide 1 through Slide 10 and 15 questioned unknown fired casings labeled with an alpha character. The participants examined and compared the 15 questioned unknown fired casings to the 10 pairs of known test fired casings, which were labeled Slide 1 through Slide 10, and were asked to determine which slides were used to fire the 15 questioned unknown fired casings.

For the durability study, each participant received 5 additional questioned casings with a new questionnaire/answer sheet. The participants examined and compared the five questioned unknown fired casings to the 10 pairs of known test fired casings that they previously received, which were labeled Slide 1 through Slide 10, and were asked to determine which slides were used to fire the five questioned unknown fired casings.

Main Analyses

The first research question asked whether firearm and tool mark examiners would be able to identify the firearms that fired the questioned casings when examining casings fired through consecutively manufactured slides. In the null hypothesis, the average error rate (as previously defined) is zero versus the alternate hypothesis in which the error rate is greater than zero. The overall average error rate was 0.000636 and the standard error was 0.006617. All analyses in the study were conducted through nonparametric methods. The Wilcoxon Signed Rank Test was used to answer the first question. With a significance level of 0.05, the p-value was 0.079, and the error rate is not significantly different from zero. Inconclusive results were not counted in the calculation of the overall average error rate.

The second research question asked whether firearm and tool mark examiners with less than 10 years of experience would reach the same conclusions as those with greater than 10 years of experience when examining casings fired through consecutively manufactured slides. Nonparametric tests on the error rate between the two populations of experience (< 10 years = 1; > 10 years = 2) were conducted. Again, inconclusive results were not counted. The Wilcoxon Rank Sum test (nonparametric test) was utilized due to the possible lack of normality. The p-value was 0.9426. The high p-value indicates that the examiners with less than 10 years of experience will not reach different conclusions than the examiners with greater than 10 years of experience. As found in **Table 1**, there was no significant difference in the error rate between the two populations.

	YRS EXP = 1	YRS EXP = 2
$\bar{X}$	0.0006536	0.00062111
s	0.00601	0.00666

Table 1: Comparison Between Years of Experience Results

The third research question asked whether firearm and tool mark examiners would be able to identify the firearms that fired the questioned casings when examining casings fired at different intervals (durability) through consecutively manufactured slides. In the null hypothesis, the average error rate for Phase 2 testing (Durability) is zero; the alternative hypothesis is that the error rate for Phase 2 testing (Durability) is greater than zero. The Wilcoxon Signed Rank

Test (nonparametric test) was utilized to determine whether the error rate was significantly higher than 0. There was a total of 114 data points. The overall average error rate was 0.0017699. The standard deviation was 0.0188, and the p-value was 0.3216. Inconclusive results were not counted in the calculation of the overall error rate.

Additional Analyses

This research study was not designed to carry out all of the below listed analyses. These analyses will serve as a guideline for future research studies. The error rates for these analyses were not significantly different from zero.

Analyses were conducted to determine if the type of microscope, lighting and/or method affected the error rate for Phase 1 testing. The Kruskal Wallis Test, which is nonparametric, was utilized. For each parameter, the null hypothesis was that the ability to identify same gun evidence would not be affected. The significance level was 0.05.

- Is the error rate different for different types of lighting? The p-value was 0.3047 .
- Is the error rate different for different microscopes? The p-value was 0.3883.
- Is there a difference in the error rate due to different methods? The p-value was 0.8297.

Analyses were conducted to determine if the type of microscope, lighting and/or method affected the error rate for Phase 2 testing. The Kruskal Wallis Test was utilized for the microscope, lighting and method. For each parameter, the null hypothesis was that the ability to identify same gun evidence would not be affected. The significance level was 0.05.

- Does the error rate depend on the lighting? The p-value was 0.9082 .
- Does the error rate depend on the type of microscopes used? The p-value was 0.8878.
- Does the error rate depend on the method being used? The p-value was 0.715.

Based on the above listed p-values for both Phase 1 and Phase 2 testing, no significant difference in error rates was observed as a function of variation in lighting, microscope or method.

Inconclusive Results

Inconclusive answers were not used to calculate the overall average error rates for Phase 1 and Phase 2 testing because they were not considered errors. According to Peterson and Markham (1995), inconclusive answers are neither incorrect nor correct and may indeed be the most appropriate response in a situation in which the sample, lab policy, and/or examiner capabilities do not permit a more definitive conclusion.

Summary of Results

The first research question asked if firearm and tool mark examiners would be able to correctly identify the firearms that fired the questioned casings when examining casings fired through consecutively manufactured slides. The dependent variable (accuracy) and the independent variable (consecutively manufactured slides) were measured by whether or not the questioned casings could be correctly identified to the consecutively manufactured slides by using individual, unique and repeatable striations/impressions (proportion of incorrect identifications). The analysis of the data revealed that the error rate was not significantly different from zero (0.000636).

The second research question asked if firearm and tool mark examiners with less than 10 years of experience would reach the same conclusions as those with greater than 10 years of experience when examining casings fired through consecutively manufactured slides. The dependent variable (accuracy) was compared against the independent variable of years of experience (knowledge and practical wisdom). The analysis of the data revealed that there were no significant differences between the two groups (less than 10 years of experience, 0.0006536, and greater than 10 years of experience, 0.00062111) and their ability to identify same gun evidence.

The third research question asked if firearm and tool mark examiners would be able to correctly identify the firearms that fired the questioned casings when examining casings fired at different intervals through consecutively manufactured slides (durability). The dependent variable (accuracy) and the independent variable (interval of firing) were measured by whether or not the questioned casings could be correctly identified to the consecutively manufactured slides by using individual, unique and repeatable striations/impressions (proportion of incorrect identifications). With a significance level of 0.05, the error rate was not significantly different from zero (0.0017699).

Demographic variables analyzed included the type of lighting, type of microscope and method. These variables were analyzed to determine if they affected the error rate.

With a significance level of 0.05, the type of lighting, type of microscope, and method did not significantly affect the error rate.

The first hypothesis states that firearm and tool mark examiners will be able to correctly identify unknown casings to the firearms that fired them when examining casings fired through consecutively manufactured slides by utilizing individual, unique and repeatable striations/impressions. The findings of this research study support this hypothesis. With a significance level of 0.05, the data revealed that the error rate was not significantly different from zero.

The second hypothesis states that the experience level of firearm and tool mark examiners will not affect identification of same gun evidence when examining casings fired through consecutively manufactured slides. The findings of this research study support this hypothesis. Based on this study, the experience level of the firearm and tool mark examiner did not affect the firearm and tool mark examiner's examination/comparison conclusions when examining casings fired through consecutively manufactured slides. With a significance level of 0.05, the analysis of the data revealed that there were no significant differences between the two groups of examiners

The third hypothesis states that firearm and tool mark examiners will be able to correctly identify unknown casings to the firearms that fired them when examining casings fired at different intervals through consecutively manufactured slides by utilizing individual, unique and repeatable striations/impressions. The findings of this research study support this hypothesis. With a significance level of 0.05, the error rate was not significantly different from zero.

The findings of this research study supports the theory in firearm and tool mark identification, that, assuming no subclass influences, each firearm/tool produces a signature of identification (striation/impression) that is unique to that firearm/tool. Through examining the individual striations/impressions, the signature can be positively identified to the firearm/tool that produced it. Such tool mark identifications are made to a practical certainty. These identifications are not absolute because it will never be possible to examine every firearm or tool in the world, a prerequisite to making absolute determinations. The conclusion that "sufficient agreement" exists between two tool marks (test and questioned) for identification means that the likelihood that another tool (firearm) could have made the questioned tool mark is so remote as to be considered a practical impossibility.

Practical impossibility currently cannot be expressed in mathematical terms. As a result of extensive empirical research and validation studies, such as this one, that have been conducted in the field of firearm and tool mark identification, as well as the cumulative results of training and casework examinations that have been either performed or peer reviewed by the examiner, an opinion can be justifiably formed that it is a practical impossibility that another firearm will be found that exhibits as much individual microscopic agreement with test tool marks as the questioned tool marks that have been identified.

For Phase 1 testing, there were a total of 3,255 questioned unknown fired casings examined by the participants. There were 3,239 correct answers, 2 incorrect answers and 14 inconclusive answers. **Table 2** illustrates the number of incorrect and inconclusive results. A total of 211 of 217 participants correctly identified same gun evidence. Two different participants reported incorrect answers. One of the two participants who reported an incorrect answer also reported one inconclusive answer. Two participants reported five inconclusive answers each. Another participant reported two inconclusive answers. Finally, one participant reported one inconclusive answer.

	Participants	**Incorrect Responses**	**Inconclusive Responses**
	n = 217		
	1	1	1
	1	1	0
	1	0	1
	1	0	2
	2	0	5
	211	0	0
Total	**217**	**2**	**14**

Table 2: Incorrect and Inconclusive Results, Phase 1

For Phase 2 testing (Durability), there were a total of 570 questioned unknown fired casings examined by the participants. There were 564 correct answers, 1 incorrect answer and 5 inconclusive answers. **Table 3** illustrates the number of incorrect and inconclusive results. A total of 112 of 114 participants correctly identified same gun evidence. One participant reported one incorrect answer. Furthermore, another participant reported five inconclusive answers.

The error rate for this research study was computed on an individual level for all participants and then averaged.

CONCLUSIONS

	Participants	Incorrect Responses	Inconclusive Responses
	n = 114		
	1	1	0
	1	0	5
	112	0	0
Total	114	1	5

Table 3: Incorrect and Inconclusive Results, Phase 2

This research study provided pertinent information relative to the forensic science community and the forensic science discipline of firearm and tool mark identification. This research study was the first investigation to utilize multiple participants (over 200) to examine fired casings from consecutively manufactured slides in order to determine an error rate. Results from this study show that firearm and tool mark examiners can accurately identify casings that were fired through consecutively manufactured slides utilizing individual (no subclass influence), unique and repeatable striations/impressions.

Consecutively manufactured slides represent a situation where the same tools and machining processes are utilized back-to-back on one slide after another. This represents a situation where the most similarity should be seen between slides. If there were ever any chance for duplication of individual marks, it would have occurred here.

The results of this research study, as well as past studies, indicate that sufficient empirical evidence exists to support the scientific foundation of firearm and tool mark identification, in which, once the specter of subclass influence is eliminated, each firearm/tool produces a signature of identification (striation/impression) that is unique to that firearm/tool. Through the examination of the individual striations/impressions, the tool mark signature can be positively identified to the firearm/tool that produced it (Ogihara et al., 1983; Shem & Striupaitis, 1983; Matty, 1984; Thompson, 1994; Hamby, 2001; Bunch & Murphy, 2003; Coody, 2003; Coffman, 2003; Vinci et al., 2005; Gouwe et al., 2008).

This research also indicates that firearm and tool mark examiners will be able to correctly identify unknown casings to the firearms that fired them when examining casings fired at different intervals through consecutively manufactured slides utilizing individual, unique and repeatable striations/impressions.

Data also revealed no significant differences in the error rate between identifications made by firearm and tool mark examiners with < 10 years of experience (0.0006536, n = 102) as compared to identifications made by examiners with > 10 years of experience (0.00062111, n = 115) when examining casings fired through consecutively manufactured slides. These results indicate that a trained firearm and tool mark examiner with two years of training, regardless of experience, will correctly identify same gun evidence.

The most significant finding in this study was the low error rate for the examination of unknown casings and identification to the firearms that fired them when examining casings fired through consecutively manufactured slides utilizing individual, unique and repeatable striations/impressions. The error rate of the participants was established by Dr. Gulati to be 0.000636 for the initial test and 0.0017699 for the durability testing. Both error rates are not significantly higher than zero.

Finally, this research study addressed concerns that were raised by the National Academy of Sciences Report (2009). The National Academy of Sciences Report questioned the repeatability and uniqueness of striations/impressions left on fired evidence used to identify same gun evidence as well as the error rate in firearms identification. Based on this research study, firearm and tool mark examiners demonstrated a very low error rate when comparing casings fired in consecutively manufactured slides.

Limitations

The researchers discovered the following limitations to this study:

- The same firing pin was not used to fire all of the known and unknown casings.
- The researchers assumed that the participants followed appropriate AFTE procedures, as listed in the AFTE Procedures Manual, FA-IV-13, Microscopic Comparison (2001).
- Each participant was administered the experimental exercise at their own crime laboratory via mail, and the researchers had no observable control.
- The researchers had to assume that each participant independently completed the experimental exercise with no outside assistance.
- The researchers had no control of the equipment that participants utilized for the experimental exercise.

- The researchers had to assume that the equipment utilized was appropriate, properly maintained and properly functioning.

- The researchers had no control over the training, skill level or experience of the participants.

- The instrument for the experimental and durability exercises was individually administered utilizing the United States Postal Service according to the email response of the participants. All eligible firearm and tool mark examiners were invited to participate.

- While the researchers personally mailed the experimental exercise to one participant per crime laboratory, that participant in turn maintained control of the exercise. The researchers had no observable control.

- The issue of accreditation was not addressed in this research study.

- The researchers had no control of the development and maintenance of standards utilized by the participants' laboratories.

- The researchers did not examine the function of individual certification in firearm and tool mark examination/identification in this research study. This information was not captured during the Phase 1 testing.

- The participants could have assumed, due to the format of the questionnaire/answer sheets and no directions to the contrary, that each set of unknowns was closed, such that each unknown casing should properly be associated with one of the test slides.

Recommendations for Future Research

Future research is needed in the forensic science community in the area of multiple consecutively manufactured slides. Considerable research has been conducted on multiple consecutively manufactured slides/breech faces (Ogihara et al., 1983; Shem & Striupaitis, 1983; Matty, 1984; Thompson, 1994; Hamby, 2001; Bunch & Murphy, 2003; Coody, 2003; Coffman, 2003; Vinci et al., 2005; Gouwe et al., 2008); however, the present research study was the first investigation to utilize multiple participants (over 200) to examine fired casings from consecutively manufactured slides in order to determine an error rate. Participants from 157 crime laboratories in 46 states plus the District of Columbia participated in this study, and additional participants from the remaining crime laboratories and states should be sought out. Future research should include a re-test of the original participants to examine repeatability of the results.

Future research should analyze the repeatability and uniqueness of striations/impressions. Additional recommendations include the following:

- Other calibers of firearms should be examined.

- Both fired bullets and casings should be examined.

- Striated tool marks should be examined utilizing the comparison methods of pattern matching and QCMS to determine if there is a difference in error rates.

- The effect of membership in professional organizations should be investigated to determine if there is an impact on results.

- The topic of accreditation should be explored to determine if accreditation of the participant's laboratory has any effect on the examination and comparison of firearm and tool mark evidence.

- Examine whether individual certification affects the outcome of the examination and comparison of firearm and tool mark evidence.

- Use an "open set" design where the participant has no expectation that all questioned tool marks should match one or more of the unknowns.

Additional research will continue to improve the scientific foundation of forensic firearm and tool mark identification through evaluation, testing and study to determine the uniqueness of striations/impressions. Furthermore, it will allow the error rates for identifications of same gun evidence to be calculated from the additional data. Fundamental research will continue to improve the understanding of the accuracy, reliability and validity of the forensic science discipline of firearm and tool mark identification.

ACKNOWLEDGMENTS

Thank you to Dr. James Hamby who permitted the researchers to adapt the Consecutively Rifled P85 Barrel Test Set Survey/Answer Sheet.

The authors also thank John Murdock and his colleagues at the Contra Costa County Sheriff's Crime Laboratory for their assistance in the editorial review of this paper, especially for the description of "practical impossibility."

Miami-Dade Police Department Crime Laboratory Firearm and Tool Mark Examiners:

- Chris Barr
- Jorge Bello
- Jill Errickson
- Yamil Garcia
- Angela Garvin
- George Hertel
- Julie Knapp
- John Mancini
- Magen Roberts
- Jami Taylor
- Tim Wilmot
- Erin Wilson

DISSEMINATION OF RESEARCH FINDINGS

NIJ 2011 Grantees Meeting Presentation, February 22, 2011, Chicago, IL.

AFTE 2011 42nd Annual Training Seminar, May 29 – June 3, 2011, Chicago, IL.

REFERENCES

[1] *Association of Firearm and Tool Mark Examiners Procedures Manual*. Version: July 9, 2001.

[2] Association of Firearm and Tool Mark Examiners Member's Page. http://afte.org/MembersArea/promanual.shtml.

[3] Association of Firearm and Tool Mark Examiners Membership Information Page. Association of Firearm and Tool Mark Examiners. http://www.afte.org/Membership/membership.htm. 16 June 2009

[4] Brundage, D. (1998). "The Identification of Consecutively Rifled Gun Barrels." AFTE Journal *30*(3), 438 – 444.

[5] Bunch, S., & Murphy, D. (2003). "A Comprehensive Validity Study for the Forensic Examination Of Cartridge Cases." AFTE Journal *35*(2), 201 – 203.

[6] Christensen L. (2004). Experiment Methodology, (9th ed.). Boston, MA: Pearson.

[7] Coody, A. (2003). "Consecutively Manufactured Ruger P-89 Slides." AFTE Journal, *35*(2), 157 – 160.

[8] Coffman, B. (2003). "Computer Numerical Control (Cnc) Production Tooling and Repeatable Characteristics on Ten Remington Model 870 Production Run Breech Bolts." AFTE Journal, *35*(1), 49 -54.

[9] Creswell, J. (2005). Educational Research: Planning, Conducting, and Evaluating Quantitative And Qualitative Research. (2nd ed.). Upper Saddle River, NJ: Pearson.

[10] Gall, M., & Borg, W. (1996). Educational Research (6th ed.). White Plains, NY: Longman.

[11] Fadul, T. (2011). "An Empirical Study to Evaluate the Repeatability and Uniqueness of Striations/Impressions Imparted on Consecutively Manufactured Glock Ebis Gun Barrels." AFTE Journal *43*(1), 37 – 44.

[12] Gouwe, J., Hamby, J., & Norris, S. (2008). "Comparison of 10,000 Consecutively Fired Cartridge Cases from a Model 22 Glock .40 S&W Caliber Semiautomatic Pistol." AFTE Journal *40*(1), 57 – 63.

[13] Green, S. (1991). "How Many Subjects Does it Take to Do a Regression Analysis." Multivariate Behavioral Research, *26*(3), 499-501. Retrieved October 13, 2008, from Academic Search Premier Database.

[14] Hamby, J. (2001). *Forensic Firearms Examination.* Unpublished doctoral thesis, University of Strathclyde, Glasgow, Scotland, U.K.

[15] Hamby, J., Brundage, D., & Thorpe, J. (2009). "The Identification of Bullets Fired from 10 Consecutively Rifled 9mm Ruger Pistol Barrels: a Research Project Involving 507 Participants from 20 Countries." AFTE Journal *41*(2), 99 – 110.

[16] Matty, W. (1984). "Raven .25 Automatic Pistol Breech Face Tool Marks." AFTE Journal *16*(3), 57 – 60.

[17] National Academy of Sciences. Strengthening Forensic Science in the United States: A Path Forward. Washington, D.C.: The National Academies Press, 2009.

[18] National Research Council (1996). The Evaluation Of Forensic DNA Evidence. National Academy Press, Washington, D C.

[19] Ogihara, Y., Kubota, M., Sandra, M., Fukuda, K., Uchiyama, T., & Hamby, J. (1983). "Comparison of 5,000 Consecutively Fired Bullets and Cartridge Cases From a 45 Caliber M1911a1 Pistol." AFTE Journal *15*(3), 127 – 140.

[20] Peterson, J., & Markham, P. (1995). "Crime Laboratory Proficiency Testing Results, 1978-1991, II: Resolving Questions of Common Origin." Journal of Forensic Sciences, *40*(6), 1009-1029.

[21] Shem, R., & Striupatis, P. (1983). "Comparison of 501 Consecutively Fired Bullets And Cartridge Cases From a 25 Caliber Raven Pistol." AFTE Journal *15*(3), 109 – 112.

[22] Thompson, E. (1994). "Phoenix Arms (Raven) Breechface Toolmarks." AFTE Journal *26*(2), 134 – 135.

[23] Vinci, F., Falamingo, R., Campobasso, C., & Bailey, J. (2005). "Morphological Study of Class Characteristics Produced by Firing 2,500 Cartridges in a .45 Caliber Semiautomatic Pistol." AFTE Journal *37*(2), 368 – 372.

Miami-Dade Police Department
Crime Laboratory
9105 NW 25th Street, Miami, Florida 33172
(305) 471-2050

Firearm & Toolmark Unit

Answer Sheet: Consecutively Manufactured Slide Test Set Test Number: ________________

Name:____________________________Job Title:________________________Date:_______________

Years Experience:_____________Years Training:__________Type of Training:____________________

Brand & Model of Microscope:______________________Type of Lighting:____________________

Do you examine other types of evidence: Yes No If Yes, what other types?_____________________
____

Do you belong to a professional or forensic organization(s)? Yes No Please list:_______________

Have you attended the FBI Specialized Techniques School? Yes No CMS Trained? Yes No

Did you use Pattern Matching or CMS for this test? ______________________________________

Please microscopically compare the known test shots from each of the 10 slides (numbered 1 through 10) with the 15 questioned casings (scribed A through Z) submitted. Indicate your conclusion(s) by circling the appropriate 'alpha' designator on the same line as the known test shots indicated. Note: This test does not have to be done all at one time, but sufficient time to adequately examine this material is necessary. Although the casings have been scribed on the body, you may elect to confirm the 'identifier' and re-scribe it. Note: This test was developed to evaluate an examiner's ability to identify fired casings based on breech face marks solely. Do not use other markings (firing pin impressions, extractor marks, ejectors marks, chamber marks, etc.) for your comparisons, as they may be misleading.

Knowns Unknowns

1. A...B...C...D...E...F...G...H...I...J...K...L...M...N...O...P...Q...R...S...T...U...V...W...X...Y...Z
2. A...B...C...D...E...F...G...H...I...J...K...L...M...N...O...P...Q...R...S...T...U...V...W...X...Y...Z
3. A...B...C...D...E...F...G...H...I...J...K...L...M...N...O...P...Q...R...S...T...U...V...W...X...Y...Z
4. A...B...C...D...E...F...G...H...I...J...K...L...M...N...O...P...Q...R...S...T...U...V...W...X...Y...Z
5. A...B...C...D...E...F...G...H...I...J...K...L...M...N...O...P...Q...R...S...T...U...V...W...X...Y...Z
6. A...B...C...D...E...F...G...H...I...J...K...L...M...N...O...P...Q...R...S...T...U...V...W...X...Y...Z
7. A...B...C...D...E...F...G...H...I...J...K...L...M...N...O...P...Q...R...S...T...U...V...W...X...Y...Z
8. A...B...C...D...E...F...G...H...I...J...K...L...M...N...O...P...Q...R...S...T...U...V...W...X...Y...Z
9. A...B...C...D...E...F...G...H...I...J...K...L...M...N...O...P...Q...R...S...T...U...V...W...X...Y...Z
10. A...B...C...D...E...F...G...H...I...J...K...L...M...N...O...P...Q...R...S...T...U...V...W...X...Y...Z

Adapted from the International Forensic Science Laboratory with the permission of Dr. James E. Hamby

Appendix A: Survey Instrument (Questionnaire/Answer Sheet)

Miami-Dade Police Department
Crime Laboratory
9105 NW 25th Street, Miami, Florida 33172
(305) 471-2050

Firearm & Toolmark Unit

Answer Sheet – Phase 2:
Consecutively Manufactured Slide Test Set

Test Number: ______________

Name: ____________________________ Date: ______________

Brand & Model of Microscope: ______________________ Type of Lighting: ______________________

AFTE Certified? Yes No ABC Certified? Yes No Other Certification? ________________

Male or Female? (Please circle one)

Did you use Pattern Matching, CMS or Both for this test? ______________________________________

You must have completed the first test set in order to participate in this 2nd phase of the research study.

Please microscopically compare the known test shots from each of the 10 slides (numbered 1 through 10) with the 5 questioned casings (scribed A through Z) submitted. Indicate your conclusion(s) by circling the appropriate 'alpha' designator on the same line as the known test shots indicated. Note: This test does not have to be done all at one time, but sufficient time to adequately examine this material is necessary. Although the casings have been scribed on the body, you may elect to confirm the 'identifier' and re-scribe it. Note: This test was developed to evaluate an examiner's ability to identify fired casings based on breech face marks solely. Do not use other markings (firing pin impressions, extractor marks, ejectors marks, chamber marks, etc.) for your comparisons, as they may be misleading.

Knowns Unknowns

1. A...B...C...D...E...F...G...H...I...J...K...L...M...N...O...P...Q...R...S...T...U...V...W...X...Y...Z
2. A...B...C...D...E...F...G...H...I...J...K...L...M...N...O...P...Q...R...S...T...U...V...W...X...Y...Z
3. A...B...C...D...E...F...G...H...I...J...K...L...M...N...O...P...Q...R...S...T...U...V...W...X...Y...Z
4. A...B...C...D...E...F...G...H...I...J...K...L...M...N...O...P...Q...R...S...T...U...V...W...X...Y...Z
5. A...B...C...D...E...F...G...H...I...J...K...L...M...N...O...P...Q...R...S...T...U...V...W...X...Y...Z
6. A...B...C...D...E...F...G...H...I...J...K...L...M...N...O...P...Q...R...S...T...U...V...W...X...Y...Z
7. A...B...C...D...E...F...G...H...I...J...K...L...M...N...O...P...Q...R...S...T...U...V...W...X...Y...Z
8. A...B...C...D...E...F...G...H...I...J...K...L...M...N...O...P...Q...R...S...T...U...V...W...X...Y...Z
9. A...B...C...D...E...F...G...H...I...J...K...L...M...N...O...P...Q...R...S...T...U...V...W...X...Y...Z
10. A...B...C...D...E...F...G...H...I...J...K...L...M...N...O...P...Q...R...S...T...U...V...W...X...Y...Z

Inconclusive: ____

Adapted from the International Forensic Science Laboratory with the permission of Dr. James E. Hamby

Appendix B: Survey Instrument (Phase 2 Questionnaire/Answer Sheet)

STURM, RUGER & CO., INC.

200 RUGER ROAD • PRESCOTT AZ 86301 U.S.A. • 928-778-6555 928-778-6633 (fax) • www.ruger.com

February 16, 2010

Gabriel A. Hernandez, M.S.
Criminalist Supervisor
Miami-Dade Police Department
Crime Laboratory Bureau
9105 NW 25th Street, Room 2159
Doral, Florida 33172-1500

Dear Mr. Hernandez,

Please accept this letter as certification that the broaching operation of the breech face on the 10 slides, marked 1 through 10, were, broached in sequential order using the same cutting tool as witnessed by my Production Control Technician Paul Kennedy and Inspector Linda Coleman.

These slides are for the Model P95PR15 pistol serial number 317-29816.

Respectfully,

[redacted] (b)(6) per EOUSA

James D. Elliott
Plant Manager
Sturm, Ruger and Co., Inc.
200 Ruger Road
Prescott, Arizona 86301
[redacted] (b)(6) per EOUSA

RUGER FIREARMS • RUGER INVESTMENT CASTING

Appendix C: Sturm Ruger Certification Letter

Manufacturing Information

The following questions were asked by the researchers and the answers were provided by Rich David and James Elliot (Private Communication, 2011), Sturm, Ruger & Company, Inc.

Question One
What type of broach was used? Was it a step broach, for example, which was fit into the breech face recess and drawn across the breech face surface removing circular milling lines and leaving parallel straight lines? Is this done by hand? A previous paper on consecutively made Ruger slides (Coody, 2008) mentions a "Barrette file" being used to create the parallel lines on P89 slides.

Answer One
It is a horizontal hydraulic step broach, which uses broach oil for a lubricant / coolant. The slide is a cast part and since it is a straight pull broach, no circular milling lines are on the breech face. The only other machining to that surface is the firing pin hole which is done prior to broaching. The broach does, however, make parallel lines in the face since the individual broach bar cutters can vary a little because of sharpening.

Question Two
After the broaching operation, is there any other final finishing done to the breech area…particularly a filing or belt sanding operation to remove burrs? Or are the slides moved onto some sort of tumbling/ sand blasting step?

Answer Two
No other files or sanders touch the breech face on a P95. The firing pin hole gets chamfered after heat treat at the end of the slide process prior to bluing. This is done manually with a tool going through the barrel hole, similar to a long, thin screwdriver with a special tip to break the sharp edge of the firing pin hole to the breech face. This prevents that sharp edge from shaving brass off the cartridge primers which could eventually cause a misfire.

Question Three
Are the slides tumbled prior to hardening? If so, is it done with ceramic beads in a bowl vibrating method?

Answer Three
The slides are only tumbled AFTER heat treat using ceramic media in a vibratory bowl.

Question Four
Are the slides sand or bead blasted prior to hardening? If so, is the breech area protected during this? If it is protected, how so? With a plastic film? Is the plastic film pitted by the beads and can the breech underneath therefore be marred slightly by this pitting?

Answer Four
The slides are only sandblasted and bead blasted AFTER heat treat. There is no protection to the breech face during this process.

Question Five
Was the broach sharpened at any time during the manufacturing of the 10 slides sent to the MDPD Crime Laboratory? Also, how often are the broaches sharpened?

Answer Five
The broach was not sharpened during the 10 piece run. They usually can go 1,000 parts between sharpening although there are a lot of conditions that could affect the life between re-sharpening.

Question Six
We noted that you mentioned that the slides are tumbled after heat treat and that they are also sandblasted and bead blasted after heat treat. What is the actual order of events after heat treat?

Answer Six
The finish process sequence is as follows: heat treat, pickle in hydrochloric acid to remove any heat treat scale, tumble, sand blast, bead blast, and bluing.

Appendix D: Manufacturing Information

Document Title: An Empirical Study to Improve the Scientific Foundation of Forensic Firearm and Tool Mark Identification Utilizing Consecutively Manufactured Glock EBIS Barrels with the Same EBIS Pattern

Author(s): Thomas G. Fadul, Jr., Gabriel A Hernandez, Erin Wilson, Stephanie Stoiloff, Sneh Gulati

Document No.: 244232

Date Received: December 2013

Award Number: 2010-DN-BX-K269

"An Empirical Study To Improve The Scientific Foundation Of Forensic Firearm And Tool Mark Identification Utilizing Consecutively Manufactured Glock EBIS Barrels With The Same EBIS Pattern"

Award Number 2010-DN-BX-K269

Final Report

Submitted By:

Miami-Dade Police Department Crime Laboratory

Thomas G. Fadul Jr., Ph.D. (MDPD)
Gabriel A. Hernandez, M.S. (MDPD)
Erin Wilson, M.F.S. (MDPD)
Stephanie Stoiloff, M.S. (MDPD)
Sneh Gulati, Ph.D. (FIU)

2013

This project was supported by Award No. 2010-DN-BX-K269 awarded by the National Institute of Justice, Office of Justice Programs, U. S. Department of Justice. The opinions, findings, and conclusions or recommendations expressed in this publication/program/exhibition are those of the authors and do not necessarily reflect those of the Department of Justice.

Abstract

This research conducted an empirical study to evaluate the reproducibility and uniqueness of striations imparted by consecutively manufactured Glock Enhanced Bullet Identification System (EBIS) barrels with the same EBIS pattern, as well as to determine the error rate for the identification of same gun evidence. The foundation of firearm and tool mark identification is that each firearm/tool produces a signature of identification (striation/impression) that is unique to that firearm/tool, and through the examination of the individual striations/impressions, the signature can be positively identified to the firearm/tool that produced it. The National Academy of Sciences Report questioned the repeatability and uniqueness of striations/impressions left on fired evidence as well as the validity and error rate in firearms identification. The Miami-Dade Police Department has been researching/evaluating Glock barrels since 1994. The Glock EBIS barrel is a polygonally rifled barrel, which has a barcode-like pattern added during the manufacturing process. Consecutively manufactured EBIS barrels with the same EBIS pattern are significant to the study because these barrels will be manufactured with the same EBIS equipment/tools and exhibit a similar pattern. Even though these barrels are consecutively made, their signatures should still be different. Test sets were assembled which included test fired bullets as well as unknowns. Participants were trained firearm and tool mark examiners throughout the United States and internationally. This empirical study established an error rate of less than 1.2 percent.

This project was supported by Award No. 2010-DN-BX-K269 awarded by the National Institute of Justice, Office of Justice Programs, U. S. Department of Justice. The opinions, findings, and conclusions or recommendations expressed in this publication/program/exhibition are those of the authors and do not necessarily reflect those of the Department of Justice.

Table of Contents

EXECUTIVE SUMMARY

In 2009, the National Academy of Sciences (NAS) released a report titled "Strengthening Forensic Science in the United States: A Path Forward." In this report, the NAS questioned the repeatability and uniqueness of striations left on fired firearms evidence as well as the validity and error rate in firearms identification. The goals of this research study were: 1) to conduct an empirical study to evaluate the repeatability and uniqueness of striations imparted by consecutively manufactured barrels; and 2) to determine the error rate for the identification of same gun evidence.

Utilizing an experimental research design, this study analyzed the repeatability and uniqueness of striations on spent projectiles/bullets fired in 10 consecutively manufactured barrels by analyzing their individual striations. One semi-automatic pistol and nine additional consecutively manufactured barrels were utilized. Consecutively manufactured barrels are significant to this study because they were manufactured with the same equipment/tools. Even though these barrels were consecutively made, their signatures should be different if there is no subclass influence. Test sets assembled included known test fired bullets from each barrel as well as unknown (questioned) bullets.

Participants were trained firearm and tool mark examiners throughout the United States and abroad. One hundred fifty test sets were distributed to laboratories in 41 states, the District of Columbia and internationally. The test sets were designed to determine an examiner's ability to correctly identify bullets fired from 10 consecutively manufactured Glock Enhanced Bullet Identification System (EBIS) barrels with the same

EBIS pattern to test fired bullets fired from the same barrels. This empirical study established an error rate of less than 1.2 percent.

The 2009 NAS Report also stated that "some forensic science disciplines are supported by little rigorous systematic research to validate the discipline's basic premises and techniques." In addition, the report stated that forensic science will be improved by collaborative opportunities "with the broader science and engineering communities." The statistical analyses of this research data was performed by a professor from the Department of Statistics at Florida International University. This collaboration with an external agency to analyze the data that was collected ensures that the statistical results are reported accurately and without bias.

This research supports the scientific foundation of forensic firearm and tool mark identification through the evaluation of the repeatability and uniqueness of striations of unknown bullets. This study provides empirical data to strengthen the foundation of firearms identification and quantifies the error rate of identification of same gun evidence from the data collected.

INTRODUCTION

Purpose, Goals and Objectives

The purpose of this research was to conduct an empirical study to evaluate the repeatability and uniqueness of striations imparted by consecutively manufactured EBIS barrels with the same EBIS pattern to spent bullets as well as to determine the error rate for the identification of same gun evidence.

Limited studies have previously been conducted with consecutively manufactured barrels as well as with the Glock EBIS barrel utilizing multiple participants. The goal was to demonstrate whether or not spent bullets could be identified to the firearm that fired them through the comparison of tool marks.

The objective of this research study was to determine if trained firearm and tool mark examiners would be able to identify spent bullets to the firearms that fired them when examining bullets fired through consecutively manufactured barrels utilizing individual, unique and repeatable striations. Also, the study presented herein evaluated the experience level of trained firearm and tool mark examiners and the effect, if any, of their experience level on the results.

Review of Relevant Literature

A review of the relevant literature found articles citing the need for additional data in the field of firearm and tool mark identification. In addition, however, a multitude of studies were also identified where one individual or a small group of individuals correctly identified bullets fired consecutively in various firearms. Some of these studies also examined the repeatability and longevity of a firearm's unique signature by comparing the test fired bullets to the questioned bullets. These research studies have demonstrated that a firearm's unique signature remains identifiable even after several thousand test fires.

Scientific Scrutiny

Saks and Koehler (1991) compared all of the forensic disciplines to DNA. The authors indicated that ballistics and tool marks have eluded the standards of scientific scrutiny. The authors discussed the lack of academic counterparts, conclusions based on unsupported data, and that only a handful of seminal research exists for all forensic disciplines besides DNA.

Saks (1993) noted that forensic science developed through criminal law. The author stated that the crime laboratory analysts lack scientific training, essentially following "recipes" (p.13). Saks discussed that crime laboratories generate little research and that few universities train the crime laboratory analysts. Saks (1994) implied that identification sciences lack statistical data to support identity. He further mentioned that the data available was not suitable for statistical analysis. Saks questioned whether firearms examiners had a reliable foundation to base their opinions.

Moran and Murdock (2002) noted the judicial system was evaluating the scientific basis for tool mark identification. The authors pointed out that the courts will be seeking demonstrative proof for identifications.

Collins (2009) indicated that forensic scientists rely upon their training and experience as their foundation for their conclusions. The author mentioned that the exact chance that two firearms will reproduce the same pattern of striations is unknown; however, "research can, in fact, quantify the margins and establish useful thresholds to help" (p. 5). Saks and Koehler (2008) noted that pattern matching lacked objective standards. They further indicated that there has been a limited amount of research regarding pattern matching.

Faigman (2010) suggested that forensic science (with the exception of nuclear DNA analysis) is not a science. He indicated that the courts allow forensic expertise that is not well validated by contemporary scientific standards. Faigman implied that practitioners are not qualified to carry out good scientific research.

Mnookin (2010) stated that the problem with forensic science is that the claims made are not supported by research. The foundation of forensic science was built on experience, not research. Mnookin is looking for more data to support examiner claims.

Ruger Consecutively Manufactured Gun Barrels

Brundage (1998) conducted an empirical study to determine whether or not firearm and tool mark examiners could properly identify bullets that were fired from consecutively manufactured Ruger gun barrels. Brundage obtained 10 consecutively manufactured 9mm Ruger firearm barrels from Sturm, Ruger & Company. The test sets were sent to 30 firearm and tool mark examiners throughout the United States. All 30 examiners returned their answer sheet; evaluation of the submitted results revealed no incorrect answers. The examiners were able to correctly distinguish the questioned bullets from multiple consecutively manufactured gun barrels. The data collected demonstrated that consecutively manufactured gun barrels differ from each other, producing different signatures. Generally speaking, this data indicated that, on a national level, firearm and tool mark examiners can identify bullets as having been fired through a particular barrel.

Hamby (2001) also conducted an empirical study to determine whether or not firearm and tool mark examiners could correctly identify bullets that were fired from consecutively manufactured gun barrels. Hamby obtained the 10 consecutively manufactured 9mm firearm barrels that were utilized in the Brundage study. The test sets were sent to 204 firearm and tool mark examiners, which included the 30 participants from Brundage's 1998 study. Hamby reported that a total of 201 examiners from several countries returned their answer sheets, and that no incorrect identifications were made.

The examiners were able to distinguish the questioned bullets from multiple consecutively manufactured gun barrels. The data collected demonstrated that consecutively manufactured gun barrels differ from each other, producing different signatures. This data also supported the hypothesis that firearm and tool mark examiners, on an international level, can identify bullets as having been fired through a particular barrel with a reasonable degree of scientific certainty.

Hamby and Brundage (2007) continued the quest of the 1998 Brundage study using the 9mm Ruger firearm barrels. A total of 438 examiners from 17 countries participated and no incorrect identifications were made. In the United States, 47 states were represented in this study. Hamby reported an error rate of .001 percent based on the data collected from all 438 participating examiners. According to Nichols (2007), "error rates have been studied and can provide consumers of the discipline with a useful guide as to the frequency with which misidentifications are reported in the community using appropriate methodologies and controls."

In 2009, Hamby, Brundage and Thorpe reported that their 10 consecutively manufactured Ruger barrel research project had a total of 507 participants from 20 countries. As of their publication in 2009, no incorrect answers were reported.

Each study described here independently tested the question as to whether markings imparted to bullets fired through consecutively manufactured barrels are reproducible and identifiable. The results of each study demonstrated that firearm and tool mark examiners can correctly identify same gun evidence. Repeated support of the same hypothesis supports the foundational concepts of firearm and tool mark examination.

Related Research

Miller (2000) conducted research to determine whether or not a trained firearm and tool mark examiner could properly identify bullets that were fired from two consecutively manufactured .44 caliber barrels. Miller found a "significant reproduction of subclass characteristics" (p. 262). Even with the amount of subclass characteristics present, Miller was still able to correctly distinguish the questioned bullets from the consecutively manufactured gun barrels.

Polygonally Rifled Barrels

Haag (1977) obtained one Heckler and Koch P9S pistol with polygonal rifling from the manufacturer for his study. Haag reported that the barrels of Heckler and Koch pistols were hammer forged, which is "a process that involves no cutting as the steel is compressed around the form" (p. 46). Haag indexed the bullets prior to test firing in order to assist with orientation for microscopic examinations. Haag reported that there were some matching striations amongst some of the bullets; however, "others revealed no positive comparison" (p. 46).

Freeman (1978) obtained three consecutively manufactured 9mm caliber Heckler and Koch polygonally rifled firearm barrels. Freeman was able to correctly distinguish the questioned bullets from the consecutively manufactured Heckler and Koch polygonally rifled firearm barrels demonstrating that consecutively manufactured gun barrels differ from each other, producing different signatures. The key limitation reported by Freeman was that one of the Heckler and Koch polygonally rifled firearm barrels used in his study did not mark as well as the other two.

Hall (1983) obtained four consecutively manufactured polygonally rifled Shilen rifle barrels. He was able to correctly distinguish the questioned bullets fired from the consecutively manufactured polygonally rifled Shilen rifle barrels demonstrating that consecutively manufactured gun barrels differ from each other, producing different signatures. Like Miller (2000), Hall (1983) noted that a subclass characteristic was present; however, it would not create a false identification.

Hocherman, Giverts and Shosani (2003) conducted a research study to determine whether or not a firearm and tool mark examiner could properly identify polygonally rifled bullets to the manufacturer of the firearm that it was fired in. Three types of polygonally rifled pistols were obtained which fit two subclass groupings of polygonal rifling profiles. The researchers created known standards (test fired bullets) and question bullets using different Glock, Jerico and Heckler and Koch pistols. Six examiners were used in this study, and they had a 65% success rate in determining the manufacturer. The researchers reported that the 65% success rate was due to a lack of training with polygonal rifling profiles.

The New York Police Department (NYPD) (1996) conducted a research study comparing bullets that were fired through polygonally rifled barrels and conventionally rifled barrels. The main purpose of the study was to determine the suitability of fired bullets for microscopic comparisons. The NYPD fired 10 cartridges through 20 Glock polygonally rifled barrels and 20 Glock conventionally rifled barrels (special order). The NYPD concluded that the ability to identify bullets that were fired through polygonally rifled barrels would be unlikely due to the barrels' inability to reproduce their signatures.

They also found that conventionally rifled barrels produced better microscopic marks for identification than polygonally rifled barrels.

Valdez (1997) examined sets of bullets fired from thirty .40 S&W Heckler and Koch polygonally rifled firearm barrels. He concluded that the difficulty of identifying these bullets was the same as cut rifling. Valdez correctly identified 28 out of 30 sets and reported that the striations appeared to be accidental and unique.

Haag (2003; 2006) introduced bore lapping, a method that utilized a grinding compound to individualize polygonally rifled barrels. He found that placing a couple of drops of a rubbing compound on the nose of a bullet that was fired in the weapon created reproducible, identifiable striations. Northcutt (2010) conducted a research project utilizing a valve-grinding compound to create reproducible, identifiable striations, which supported Haag's findings. In 2009, L. Haag, M. Haag, Garrett, Knell and Patel reported that bore lapping produced identifiable striations.

The Miami Barrel – Enhanced Bullet Identification System (EBIS)

Carr and Fadul (1997) conducted a study to determine whether or not a trained firearm and tool mark examiner could readily identify bullets that were fired from 22 different pistols and five Glock barrels marked with the electronic spark reduction method. Three firearm and tool mark examiners participated in this study. This study found that all of the weapons except Glock and Heckler and Koch marked the bullets in a readily identifiable state. The standard Glock barrels and the five Glock barrels marked with the electronic spark reduction method were listed as not readily identifiable. The inability to readily identify bullets fired in these Glock barrels began the evolution of what would become known as the Miami Barrel.

Fadul and Nunez (2003) found that Glock pulled a single cutter through their polygonally rifled barrel to create the Miami Barrel; this cutter created a subclass characteristic. A study was then conducted on the Miami Barrel to determine whether or not Glock could reproduce identifiable striations that could be readily identifiable. Fadul and Nunez used 22 Miami Barrels manufactured by Glock. Nine firearm and tool mark examiners participated in this study. All nine examiners concluded that the new Miami Barrel was not readily identifiable.

Glock then created a new version of the single cutter used in the Fadul and Nunez (2003) study. Glock called the new cutter the EBIS; however, the barrel itself is still known as the Miami Barrel. Fadul and Nunez (2006) conducted a follow-up study with three new Miami Barrels to determine whether or not these barrels could reproduce identifiable striations that could be readily identifiable. Fadul and Nunez concluded that the new Miami Barrel manufactured with the EBIS was readily identifiable. The key limitation to this study, however, was that only three barrels were examined and a concern was expressed regarding subclass characteristics. The greater concern was that an examiner who is not familiar with these markings will rely on the subclass characteristics alone for a positive identification.

Chin and Sampson (2007) followed up the Fadul and Nunez (2006) study on the Miami Barrel manufactured by Glock to determine whether or not the EBIS reproduced identifiable striations that would allow questioned fired bullets to be identified to known standards. The researchers used four Miami Barrels manufactured by Glock, which incorporated the EBIS. The questioned bullets and known standards were correctly identified. They expressed the same concern as Fadul and Nunez (2006) regarding the subclass characteristics.

Martinez (2008) conducted a study to test the durability of the Miami/EBIS barrel to determine if the EBIS barcode reproduced identifiable striations that would allow questioned fired bullets to be identified to known standards. Martinez used 51 Glock pistols which incorporated the EBIS barrel. A three year window existed between the initial test firing and the final test firing for this research study. Each pistol had at least 250 rounds fired through the barrel, and no more than 10,000 maximum. The Martinez (2008) study reported that 29% of the participants (8 out of 28) with 5 to 10 years of experience reported via survey that there were not enough individual characteristics present to conclude an identification and/or elimination. Additionally, 14% of the participants (4 out of 28) with 5 to 10 years of experience reported identifications and the ability to eliminate. Martinez believed that the identifications were made utilizing the process of elimination.

Fadul (2011) conducted a study utilizing 10 consecutively manufactured Glock Miami/EBIS barrels to further explore the capability of identifying bullets fired through the Glock Miami/EBIS barrels. On a voluntary basis, 150 test sets were created and distributed to laboratories in forty-four states and nine countries. The test set was designed to determine an examiner's ability to correctly identify questioned bullets fired from 10 consecutively manufactured Miami/EBIS barrels to test fired bullets from the same barrels. Fadul reported that 183 participants made 2,734 correct identifications and that 7 participants accounted for 11 errors. Fadul reported an error rate of 0.4 percent.

Fadul also reported that the 10 barrels utilized did not have the same EBIS barcode pattern.

Limited studies have been conducted to identify bullets from consecutively manufactured barrels where the number of participants in the study reached or exceeded one hundred. A larger sample size will provide a more reliable estimate of the true error rate for the identification of same gun evidence by firearm and tool mark examiners.

No study has been conducted to identify bullets from consecutively manufactured Glock EBIS barrels with the same EBIS patterns. The EBIS barrel is a polygonally rifled barrel, which has a barcode-like pattern added to it during the manufacturing process. Previous research included consecutively manufactured EBIS barrels; however, some of the barcode-like patterns were different.

Utilizing barrels with the same EBIS pattern and a larger sample size will provide a more precise error rate calculation for the identification of same gun evidence by firearm and tool mark examiners. Additionally, it will document the reliability, reproducibility, and the individuality of the EBIS barrels.

RESEARCH DESIGN AND METHODS

This study utilized an experimental research design (Christensen, 2004; Creswell, 2005), and was conducted in a crime laboratory setting. Participants compared questioned bullets to known standards that were fired in 10 consecutively manufactured barrels in order to determine whether or not the consecutively manufactured barrels produced different signatures, each with unique striations (tool marks). This research study also established an error rate for the identification of same gun evidence.

Quantitative data was utilized to determine if trained examiners could correctly identify questioned bullets test fired in multiple consecutively manufactured barrels. Additionally, the number of years of experience of the examiners was recorded. This data was used to draw conclusions for the following: 1) whether or not trained firearm and tool mark examiners can identify same gun evidence fired though consecutively manufactured EBIS barrels; and 2) whether years of experience impacts correct identifications. Questionnaire/answer sheets were utilized to collect the quantitative data (see Appendix A).

The proposed outcome in this section is presented with the intention that the findings will be able to answer the research questions.

Research Questions

Q1. Will trained firearm and tool mark examiners be able to correctly identify the firearms that fired the questioned bullets when examining bullets fired through consecutively manufactured barrels with the same EBIS pattern?

Q2. Will firearm and tool mark examiners with less than 10 years of experience reach the same conclusions than those with greater than 10 years of experience when examining bullets fired through consecutively manufactured barrels with the same EBIS pattern?

Research Hypotheses

H1. Trained firearm and tool mark examiners will be able to correctly identify unknown bullets to the firearms that fired them when examining bullets fired through consecutively manufactured barrels with the same EBIS pattern utilizing individual, unique and repeatable striations.

H2. The experience level of firearm and tool mark examiners will not affect identification results when examining bullets fired through consecutively manufactured barrels with the same EBIS pattern.

There is one dependent variable that was examined in this study. The dependent variable is inaccuracy (proportion of incorrect identifications), which was measured by whether or not the questioned bullets could be correctly identified to the consecutively manufactured barrels by using individual, unique and repeatable striations.

There are several independent variables in this study, such as the consecutively manufactured barrels, interval of firing and experience of the examiner. For Q1 and H1, the researchers were interested in studying the effect of the consecutively manufactured barrels on the ability to identify same gun evidence. For Q2 and H2, the researchers were interested in studying the effect of the independent variable of experience, the knowledge and practical wisdom gained through study, observation, experimentation and casework, on the ability to identify same gun evidence. Extraneous variables were controlled as much as possible by utilizing laboratory settings.

For Q1 and Q2, the dependent variable was measured through the average error rate as determined from the *Consecutively Rifled EBIS-2 Test Set* questionnaire/answer sheet using a 1 to 10 point system (1 point for each correct answer, with a maximum point value of 10).

Two hypotheses, H1 and H2, were tested in this study. For H1, the dependent and independent variables measured whether or not consecutively manufactured barrels with the same EBIS pattern produced individual, unique and repeatable striations based on each participant's results. If the tool mark signatures from each of the ten consecutively manufactured barrels could be distinguished from one another by the participants, this would establish that there is no subclass tool mark influence present from the manufacturing process used to form the barrels that affects the ability to identify same gun evidence.

For H2, the dependent and independent variables measured whether or not the years of experience of the participants affected the ability of the examiners to identify same gun evidence based on each participant's results.

Target Population

In this study, the target population represented a subset of the forensic science community, more specifically, firearm and tool mark examiners employed by a national or international law enforcement agency (crime laboratory), or like agency. The Miami-Dade Police Department (MDPD) Crime Laboratory (CL) utilized the membership list for the Association of Firearm and Tool Mark Examiners (AFTE); one of the privileges of membership is access to the membership list. Twelve members of this association currently work in the MDPD CL.

Membership in AFTE is limited to individuals with suitable education, training, and experience in the examination of firearms and/or tool marks. For purposes of this membership, a practicing firearm and/or tool mark examiner is defined by AFTE (2009) as: "(1) An individual who derives a substantial portion of his livelihood from the examination, identification, and evaluation of firearms and related materials and/or tool marks; and (2) An individual whose present livelihood is a direct result of the knowledge and experience gained from the examination, identification, and evaluation of firearms and related materials and/or tool marks."

Every firearms examiner in the United States and internationally who is a member of AFTE had an equal opportunity to be included in this study. Each AFTE member was contacted by the MDPD CL via email inviting them to participate in this study, which included completing demographic questions and participation in an experimental exercise. The test sets utilized in this study were similar to the work that the participants perform on a routine daily basis.

Eligibility-Inclusion Criteria

Participants were required to be firearm and tool mark examiners employed by a national or international law enforcement agency (crime laboratory), or like agency, and must have completed a two year training program. Independent examiners who retired from a qualifying agency were also eligible to participate in this study.

Accessible Population

Accessibility was limited to firearm and tool mark examiners for whom the MDPD CL was able to obtain email addresses by querying the membership list for AFTE. AFTE members with email addresses listed account for 92% of the 2010 AFTE Roster.

Sampling Plan and Setting

The sampling plan for this study utilized an abstract population. Every eligible firearm and tool mark examiner in the United States and internationally who is a member of AFTE, with a functional email address, was invited to participate in this research. The AFTE list identified participants that met the MDPD CL requirements as stated in Eligibility-Inclusion Criteria. The accessible population included approximately 800 firearm and tool mark examiners in the United States and internationally.

To ensure confidentiality, the researchers at the MDPD CL invited firearm and tool mark examiners to participate via email. The survey and test (see Appendix A) were conducted by each participant independently, which strengthened the study's validity (Gall & Borg, 1996).

Instrumentation

This study utilized two methods of instrumentation: a questionnaire that included the participant's demographics, as well as an answer sheet for an experimental exercise. The questionnaire took less than ten minutes to complete. The experimental exercise took approximately two to four hours to complete. These approximate times were based on personal communication and observation of participants from the MDPD CL.

The experimental exercise was originally utilized by Brundage (1998), redesigned by Hamby (2001) and adapted by Fadul (2011). Over 600 firearm and tool mark examiners have used this instrument (Hamby & Brundage, 2007; Fadul, 2011). The researchers at the MDPD CL modified this instrument by adding the number of years of experience for each firearm and tool mark examiner. The categories of ASCLD/LAB accreditation, AFTE certification, Miami/EBIS barrel and gender were added. Furthermore, the researchers at the MDPD CL added a category for pattern matching and quantitative consecutive matching striations (QCMS). Pattern matching and QCMS are two forms of striated tool mark examination/assessment processes used in the field of firearm and tool mark examination. The categories of "Inconclusive" and "Elimination" were added for the experimental exercise. Additionally, a field was added for "Other Results/Comments."

Data Collection Methods

The researchers did the following:

1. Received National Institute of Justice (NIJ) approval.
2. Sent email to the AFTE membership. Participation was voluntary.
3. Traveled to Glock, Inc. factory, observed the manufacturing process of the EBIS barrel, and recorded the data matrix code for each barrel (for tracking).
4. Obtained 10 EBIS barrels from Glock, labeled 1 through 10.
5. Obtained 1 pistol from the MDPD Firearms Reference Collection.
6. Obtained 9mm Federal cartridges (ammunition/bullets).
7. Utilized the MDPD CL water tank for test firing and retrieval of the bullets.
8. Placed each barrel one at a time in the pistol.
9. Loaded the pistol with five cartridges.
10. Fired the pistol into the shoot tank.
11. Fired five bullets through each barrel to create one test set. (This was repeated 150 times per barrel).
12. Used secure, properly labeled containers to keep each group of five bullets separated.
13. Labeled two of the five bullets with the number of the barrel in which they were fired (1 through 8) to create the test fired bullets (known standards). These known standards were placed in a labeled coin envelope.
14. Labeled remaining three bullets with an alpha character designated by the researchers at the MDPD CL to represent the unknown (questioned) bullets. Different alpha characters were assigned to each barrel.

15. Selected one unknown bullet from barrels 1, 2, 3, 5, 6, 7 and 8 randomly from the containers and placed it in a labeled coin envelope. No unknown bullets from barrel 4 were provided to the participants.
16. Selected one additional unknown bullet from barrel 3 and placed it in a labeled coin envelope.
17. Selected two unknown bullets from barrel 9 to complete the test set of 10 unknown bullets. These two bullets were each placed in a labeled coin envelope. No test fired bullets (known standards) from barrel 9 were provided to the participants.
18. No test fired bullets (known standards) or unknown (questioned) bullets from barrel 10 were provided to the participants.
19. Created 150 test sets and placed each test set in a medium manila envelope.
20. The researchers microscopically examined every 10th set to ensure that the bullets were comparable and identifiable.
21. Mailed test sets to respondents. Each respondent received one test packet through the mail which included the following:
    - One questionnaire/answer sheet
    - 10 unknown bullets
    - 8 sets of test fired bullets (known standards) that were fired through 8 consecutively manufactured barrels.
22. Instructed the participants via the questionnaire/answer sheet to compare the questioned bullets to the known standards, and to place their answers on the questionnaire/answer sheet.

- The participants were also asked to complete the questions that were on the questionnaire/answer sheet.
- The instructions directed the participants to return the questionnaire/answer sheet via mail or fax.

23. Conducted the data collection process for 26 weeks.
24. Utilized an Excel spreadsheet to record and analyze the data collected.
25. Submitted the data to a professor from the Department of Statistics at Florida International University for statistical analyses.

<u>Data Coding</u>

Each participant was assigned a number from 1 to end. There were 10 variables (unknown bullets) which were designated with an alpha character and coded as correct (1), incorrect (2), or inconclusive (3A – 3D; four types are detailed in Summary of Results). The overall correct number was coded 1 through 10 based on the correct number of identifications. Pattern matching was coded as 1, QCMS was coded as 2, utilization of both methods was coded as 3 and no answer was coded as 4. The type of microscope was coded 1 for Leica, 2 for Leeds, 3 for other, or 4 for no answer. Type of lighting was coded 1 for fluorescent, 2 for fiber optic, 3 for LED, 4 for other, or 5 for no answer. Years of experience was coded 1 for less than 10 years and coded 2 for greater than 10 years. Laboratory ASCLD/LAB Accreditation was coded 1 for "yes," 2 for "no" and 3 for no answer. Other accreditations were captured on the questionnaire but were not coded. AFTE certification was coded 1 for "yes," 2 for "no" and 3 for no answer. ABC certification was coded 1 for "yes," 2 for "no" and 3 for no answer. Have you ever encountered the Miami or EBIS barrel in your casework was coded 1 for "yes," 2 for "no" and 3 for no answer. Did you participate in the first Miami Barrel/EBIS study was coded 1 for "yes," 2 for "no" and 3 for no answer. The number of years of training was coded 1 for 2 years or more and 2 for < 2 years. The type of training was coded into 4 groups: 1 for in-house/structured, 2 for National Firearms Examiner Academy, 3 for other, and 4 for no answer. Individuals trained in QCMS were coded 1 for "yes," 2 for "no" and 3 for no answer (see Appendix A).

Descriptive Analysis

Descriptive analysis was used to describe the participants. Descriptive analysis included years of experience, method used (pattern matching/QCMS), accreditation, certification, as well as the type of microscope and lighting used.

Data Analysis Methods

Simple descriptive scores were used to analyze all variables. Statistical analysis was performed utilizing a statistical program, S-PLUS, to answer the two research questions. An independent statistician performed the data analyses.

Definitions

For this research study, the following definitions for confidence interval and error rate will apply.

Confidence Interval

A confidence interval is a range of values used to estimate the true value of a population factor. This study utilized a 95% confidence interval. Typically a 95% confidence interval is computed reflecting the probability that in 95% of the samples tested, the interval should contain the true value of the population factor (Butler, 2005).

Error Rate

An error rate is a calculated value that represents the comparison of the number of wrong responses with the total number of responses. The error rate for each participant was defined as the proportion of questions answered incorrectly over their total number of responses. For example, if a participant answered 3 out of the 10 questions incorrectly, their error rate is 0.3.

An average error rate is calculated by dividing the sum of the error rates per respondent by the total number of respondents. For example, if there are four participants and two of the participants answered 1 out of 10 questions incorrectly, one of the participants answered 3 out of 10 incorrectly, and one answered none of the 10 questions incorrectly, then the average error rate would be calculated as follows: $[(2 \times 0.1)+(1 \times 0.3)+(1 \times 0.0)] / 4 = 0.13$. An average error rate calculation was used for this study. An average error rate calculation was used by the researchers because it is illustrative of the error rate across all participants rather than solely based on number of responses.

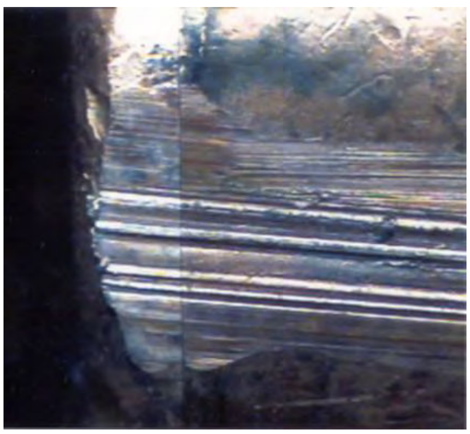
Manufacturing Process of EBIS Barrels

Standard issue Glock barrels are polygonally rifled. EBIS barrels are essentially standard issue, polygonally rifled Glock barrels with an added manufacturing step. All Glock barrels are drilled, reamed, honed and cold hammer forged on a mandrel. The mandrel is a negative of the polygonal rifling profile. The cold hammer forging operation marks the end of the machining process for a standard issue Glock barrel.

To create an EBIS barrel, a standard Glock polygonal barrel is further processed after the cold hammer forging operation. This additional machine imparts a barcode-like pattern on the surface of the lands within the interior of the barrel. The machine does this by inserting a rod outfitted with two carbide steel cutting wheels situated 180˚ from each other (Photograph 1). The rod can make up to fifteen passes for a barrel with six lands and grooves. Each land can have up to five channels cut into its surface. Because there are two cutting wheels situated 180˚ from each other, opposite lands are imparted with the same barcode-like profile. Photograph 2 illustrates the EBIS pattern on two fired bullets.

The barrels are placed on a holder to guarantee the start position and the rod is inserted from the chamber end. The rod tool with the cutting wheels remains stationary while the barrel turns. Channels are cut on the surface of opposite lands from chamber to muzzle as the barrel rotates around the tool following the rate of twist of the rifling. After one pass is made, the barrel rotates to a groove and the tool is extracted. The tool then positions itself for the next pass. According to Glock personnel, the tool does not go down the surface of a land perfectly straight during rotation, and they have no control of this action. Each barrel is inspected with a bore scope to ensure that the barcode pattern was imparted. The EBIS barcode pattern is imparted prior to hardening. Five hundred barrels can be imparted with an EBIS barcode pattern before the carbide steel cutting wheels have to be changed.

Photograph 1: EBIS tool with carbide steel cutting wheel.



Photograph 2: Photomicrograph of the EBIS pattern on two fired bullets.

<u>Internal Validity Strengths</u>

- The quantitative data was internally valid due to the procedures set forth to assemble the tests.
- All the test materials were assembled in a crime laboratory setting.
- All unknown bullets and known standard bullets were labeled with a number (known standard) or letter (unknown bullets).
- Secure containers were used to keep the unknown bullets separated into groups.

- The researchers at the MDPD CL microscopically examined every 10th test set to ensure that the bullets were comparable and identifiable.
- The questionnaire/answer sheet used has been documented in previous studies, and the sheet is a standardized format.

Internal Validity Weaknesses

- The validity of this study was dependent upon the accuracy of the assembly of the tests.
- Communication between participants could have threatened the internal validity.
- The possibility exists that the unknown bullets and known standards failed to mark clearly. Since every set was not microscopically examined to ensure that the bullets were comparable and identifiable, some sets may have contained bullets that were not suitable for identification.

External Validity Strengths

- The external validity strength of this research project was that all testing was conducted in a crime laboratory setting.
- Participants utilized a comparison microscope.
- The participants were trained firearm and tool mark examiners.
- The training and experience of the participants strengthened the external validity.
- The number of participants exceeded the minimum sample size needed to be statistically significant.

External Validity Weaknesses

- The researchers assumed that the participants followed appropriate AFTE procedures, as listed in the AFTE Procedures Manual, FA-IV-13, Microscopic Comparison (2001).

- The researchers had no control over the equipment used by the participants.
- The training and skill level as well as the experience of the participants could have been an external weakness.
- The researchers had to assume that a participant's laboratory policy could allow for eliminations based on individual characteristics instead of an inconclusive result.

RESULTS

In this section, the examination of research questions, hypotheses testing, and other findings related to this study were analyzed to evaluate the repeatability and uniqueness of striations imparted to consecutively manufactured barrels as well as to determine the error rate for the identification of same gun evidence. This experimental exercise was designed to measure accuracy. The participant's ability to perform the experimental exercises was evaluated against reported data and demographic characteristics.

For participant performance relating to accuracy and methods utilized, a mass email was sent out to the membership of the AFTE. A total of 201 examiners representing 125 crime laboratories in 41 states, the District of Columbia, and 4 international countries completed the *Consecutively Rifled EBIS-2 Test Set* questionnaire/answer sheet. Thirty-six of the 201 participants did not meet the two year training requirement for this study. This resulted in a data-producing sample of 165 participants.

The firearm and tool mark examiners that responded to the *Consecutively Rifled EBIS-2 Test Set* questionnaire/answer sheet represented 82% of the states in the United States that conduct firearm and tool mark examinations.

The questionnaire/answer sheet utilized for this study allowed the participants to record their answer by circling the appropriate known test fired bullet sets designated by a numerical number 1 – 8 on the same line as the alpha designator of the unknown bullet. The questionnaire/answer sheet also allowed the participants to record inconclusive and/or elimination. Additionally, the participants were provided a field for other results/comments.

The statistician utilized the statistical analysis program S-PLUS for this study. Nonparametric tests, namely the Wilcoxon Signed Rank and the Wilcoxon Rank Sum tests were used for the analysis. These tests were also followed up by procedures based on large sample approximations to the distribution of the average error rate(s). The Wilcoxon Signed Rank test is a nonparametric alternative to the paired Student's t-test while the Wilcoxon Rank Sum test is used for comparing two independent samples. The tests are used when sample populations cannot be assumed to follow distributional assumptions. Quite often these tests are based on ranks. As an example, when comparing two independent pools of data from two different populations (i.e. individuals having greater than 10 years of experience versus individuals having less than 10 years of experience), one would first combine the two pools of data and rank their tested values (number of incorrect responses, for example) from the lowest to the highest. The lowest observation gets rank 1, the next one rank 2, etc. After ranking the combined pool, one would then separate the data back to their original population and sum up the ranks of each data set. If the results from the two populations are roughly similar, there should be no significant difference in the sum of the ranks (adjusted for sample size). A difference between the sums of the ranks would indicate that the participants from one population performed differently than the other population.

In 2009, the National Academy of Sciences issued a report entitled "Strengthening Forensic Science in the United States: A Path Forward." This report stated that "some forensic science disciplines are supported by little rigorous systematic research to validate the discipline's basic premises and techniques" (p. S-16). In addition, the report stated that forensic sciences will be improved by collaborative opportunities "with the broader science and engineering communities" (p. S-16). The statistical analyses of this research data were performed by Dr. Sneh Gulati, a professor from the Department of Statistics at Florida International University. This collaboration with an external agency to analyze the data that is collected ensures that the statistical results are reported accurately and without bias.

Instrument Parameters

Each participant received a total of eight pairs of known test fired bullets labeled Barrel 1 through Barrel 8 and ten unknown fired bullets labeled with an alpha character. The participants examined and compared the ten unknown fired bullets to the eight pairs of known test fired bullets, and were asked to determine which barrels were used to fire the ten unknown fired bullets.

The researchers utilized an "open set" design where the participants had no expectation that all unknown tool marks should match one or more of the unknowns, as illustrated in Table 1.

Table 1

Open Set Design

Barrel #	Known Standards (Test Fired Bullets)	Unknown Bullets (Questioned)
1	Provided	C
2	Provided	H
3	Provided	A, F
4	Provided	None
5	Provided	D
6	Provided	I
7	Provided	E
8	Provided	B
9	Not Provided	G, J
10	Not Provided	None

Main Analyses

The first research question asked whether trained firearm and tool mark examiners would be able to identify the firearms that fired the unknown bullets when examining bullets fired through consecutively manufactured barrels. Answering the first research question is equivalent to testing whether the average error rate is zero against the alternate that the average error rate is greater than zero. The overall average error rate was [(156 x 0.0) + (6 x 0.1) + (3 x 0.2)] / 165 = 0.007 and the standard deviation was 0.032. All analyses in the study were conducted through nonparametric methods. The Wilcoxon Signed Rank Test was used to answer the first question. With a significance level of 0.05, the p-value was 0.0027, which indicates that the average error rate is significantly different from zero. A 95% confidence interval for the average error rate, based on the large sample distribution of the sample average error rate, is between 0.002 and 0.012. Using a confidence interval of 95%, the error rate is no more than 0.012, or

1.2%. Inconclusive results were not counted in the calculation of the overall average error rate.

The second research question asked whether trained firearm and tool mark examiners with less than 10 years of experience would reach the same conclusions as those with greater than 10 years of experience when examining bullets fired through consecutively manufactured barrels. Nonparametric tests on the error rate between the two populations of experience (< 10 years = 1; > 10 years = 2) were conducted. Again, inconclusive results were not counted. The Wilcoxon Rank Sum test (nonparametric test) was utilized due to the possible lack of normality. The p-value was 0.9735 with a significance level of 0.05. The high p-value indicates that the examiners with less than 10 years of experience will not reach different conclusions than the examiners with greater than 10 years of experience. As found in Table 2, there was no significant difference in the error rate between the two populations.

Table 2

Comparison Of Error Rates Based On Years Of Experience

	YRS EXP = 1	YRS EXP = 2
$\bar{X}$	0.008	0.007
s	0.036	0.029

<u>Additional Analyses</u>

Certification and accreditation were evaluated to determine if they affected the error rate. Thirty-five of the 165 participants reported that they were AFTE Certified. Three of these 35 AFTE certified participants reported a total of four errors, resulting in an error rate of 0.011 for AFTE Certified participants. One hundred forty-five participants reported that their laboratory was accredited (134 ASCLD/LAB with 133 from the United States), 4 - FQS (all from the United States), 7 - AFFSAB and NATA. Seven of these 145 participants reported nine errors, resulting in error rates of 0.006 (overall) and 0.007 (United States) for participants working in an accredited laboratory. See Table 3 for certification/accreditation acronym description.

The method (QCMS and pattern matching) utilized by the participants was also evaluated. Thirty-six of the 165 participants reported that they used both QCMS and pattern matching. Two of the 36 participants reported three errors, resulting in an error rate of 0.008 for participants utilizing both QCMS and pattern matching. One hundred fifteen of the 165 participants reported that they used only pattern matching. Seven of the 115 participants reported a total of nine errors, resulting in an error rate of 0.008 for participants utilizing only pattern matching. No participant reported using QCMS by itself. Fourteen of the 165 participants did not list the method utilized.

Table 3

Certification/Accreditation Acronyms

ABC	American Board of Criminalistics
AFFSAB	Australasian Forensic Field Sciences Accreditation Board
AFTE	Association of Firearm and Tool Mark Examiners
ASCLD/LAB	American Society of Crime Laboratory Directors / Laboratory Accreditation Board
FQS	Forensic Quality Services
NATA	National Association of Testing Authorities (Australia)

Inconclusive Results

Inconclusive responses were not used to calculate the overall average error rates for this research because they were not considered errors. According to Peterson and Markham (1995), inconclusive responses are neither incorrect nor correct and may indeed be the most appropriate response in a situation in which the sample, lab policy, and/or examiner capabilities do not permit a more definitive conclusion.

Summary of Results

The first research question asked if trained firearm and tool mark examiners would be able to correctly identify the firearms that fired the unknown bullets when examining bullets fired through consecutively manufactured barrels with the same EBIS pattern. The dependent variable (inaccuracy) and the independent variable (consecutively manufactured barrels) were measured by whether or not the unknown bullets could be correctly identified to the consecutively manufactured barrels by using individual, unique and repeatable striations (proportion of incorrect identifications). The analysis of the data revealed that the error rate was significantly greater than zero, albeit less than 1.2% based on a 95% confidence level.

The second research question asked if trained firearm and tool mark examiners with less than 10 years of experience would reach the same conclusions as those with greater than 10 years of experience when examining bullets fired through consecutively manufactured barrels with the same EBIS pattern. The dependent variable (inaccuracy) was compared against the independent variable of years of experience. The analysis of the data revealed that there were no significant differences between the two groups and their ability to identify same gun evidence.

Demographic variables analyzed included certification, accreditation and method used for examination. These demographics were evaluated to determine if they affected the error rate. With a significance level of 0.05, certification, accreditation and method did not significantly affect the error rate. Note: These results should be interpreted with caution due to the large number of zero error rates. Zero error rates can lead to a large number of ties in the non-parametric test which can lead to misleading results for small samples. Since the current study has a large sample size, however, the results should not be affected.

The first hypothesis states that trained firearm and tool mark examiners will be able to correctly identify unknown bullets to the firearms that fired them when examining bullets fired through consecutively manufactured barrels with the same EBIS pattern by utilizing individual, unique and repeatable striations. The findings of this research study support the hypothesis that trained firearm and tool mark examiners can correctly identify same gun evidence with an average error rate of 0.007 (0.7%). With a significance level of 0.05, the p-value was 0.0027, which indicates that the average error rate is significantly different from zero. A 95% confidence interval for the average error rate, based on the large sample distribution of the sample average error rate, is between 0.002 and 0.012. Using a confidence interval of 95%, the error rate is no more than 0.012, or 1.2%.

The second hypothesis states that the experience level of firearm and tool mark examiners will not affect identification of same gun evidence when examining bullets fired through consecutively manufactured barrels with the same EBIS pattern. The findings of this research study support this hypothesis. Based on this study, the experience level of the firearm and tool mark examiner did not affect the firearm and tool mark examiner's examination/comparison conclusions when examining bullets fired through consecutively manufactured barrels. With a significance level of 0.05, the analysis of the data revealed that there were no significant differences between the two groups of examiners.

The findings of this research study support the theory in firearm and tool mark identification that, assuming no subclass influences, each firearm/tool produces a signature of identification (striation/impression) that is unique to that firearm/tool. Through examination of the individual striations/impressions, the signature can be positively identified to the firearm/tool that produced it. Such tool mark identifications are made to a practical certainty. These identifications are not absolute because it will never be possible to examine every firearm or tool in the world, a prerequisite to making absolute determinations. The conclusion that "sufficient agreement" exists between two tool marks (test and questioned) for identification means that the likelihood that another tool (firearm) could have made the questioned tool mark is so remote as to be considered a practical impossibility.

Practical impossibility currently cannot be expressed in mathematical terms. As a result of extensive empirical research and validation studies such as this one that have been conducted in the field of firearm and tool mark identification, as well as the cumulative results of training and casework examinations that have been either performed or peer reviewed by a trained firearm and tool mark examiner, an opinion can be justifiably formed that it is a practical impossibility that another firearm will be found that exhibits as much individual microscopic agreement with test tool marks as the questioned tool marks that have been identified.

There were a total of 1,650 unknown fired bullets examined by the participants. There were 1,496 correct answers, 12 incorrect answers and 142 inconclusive answers. The 12 incorrect answers were made by 9 participants. Two of the participants identified unknown bullets G and J (no known standard provided) to barrel 4 (no unknown bullets provided), which accounted for 4 of the 12 errors. Additionally, three participants from the same laboratory made the same error in misidentifying questioned bullet B. Due to these results, the researchers recalled this test set for examination. There were no packaging errors and bullet B was identifiable.

Sixty-nine participants had inconclusive results for unknown bullets G and J which accounted for 138 of the 142 inconclusive results. Standards from the barrel that fired unknown bullets G and J were not provided to the participants. The following illustrates the definition of an inconclusive answer for this research study. Inconclusive was coded as (3). A further breakdown of inconclusive responses is detailed here:

- **Inconclusive** response (3A): This designation means that the correct barrel number was not circled, and the Inconclusive answer was circled instead. This inconclusive response could have been selected for any of the unknowns. When specifically relating to Unknowns G and J, the Inconclusive answer was circled, and the test taker did NOT identify Unknowns G and J as having come from one firearm (22 participants, 42 of 46 inconclusive responses were for G and J).
- **Inconclusive** response (3B): <u>Relating only to Unknowns G and J</u>. The Inconclusive answer was circled, and the test taker DID identify Unknowns G and J as having come from one firearm (39 participants, 78 inconclusive responses).

- **Inconclusive** response (3C): Relating only to Unknowns G and J. No answer was circled, but the test taker DID identify Unknowns G and J as having come from one firearm (2 participants, 4 inconclusive responses).
- **Inconclusive** response (3D): Relating only to Unknowns G and J. The Elimination answer was circled, but the test taker DID NOT identify Unknowns G and J as having come from one firearm (7 participants, 14 inconclusive responses).

Table 4 illustrates the number of incorrect and inconclusive results.

Table 4

Incorrect and Inconclusive Results: Unknown (Questioned) Bullets

	Unknown Bullets (Questioned)	**Incorrect Responses**	**Inconclusive Responses**
n = 165			
	A	0	2
	B	4	0
	C	0	0
	D	0	0
	E	2	1
	F	1	1
	G	2	69
	H	1	0
	I	0	0
	J	2	69
Total	**10**	**12**	**142**

The error rate for this research study was computed on an individual level for each participant and then averaged across all participants.

CONCLUSIONS

This research study provides relevant information to the forensic science community and to the forensic science discipline of firearm and tool mark identification. This research study utilized multiple participants (n = 165) to examine fired bullets from consecutively manufactured barrels in order to determine an error rate for identification of same gun evidence. The results of this study support the hypothesis that trained firearm and tool mark examiners can identify fired bullets to the correct consecutively manufactured barrel utilizing individual (no subclass influence), unique and repeatable striations.

Consecutively manufactured barrels represent the best possibility for the production of two firearms that could produce non-distinguishable markings since the same tools and machining processes are utilized back-to-back on one barrel after another. This process thus represents a situation where the most similarity should be seen between barrels. If there were ever any chance for duplication of individual marks, it would occur during the manufacture of consecutively manufactured barrels. The results of this research study, as well as past studies, indicate that sufficient empirical evidence exists to support the scientific foundation of firearm and tool mark identification. Once the specter of subclass influence is eliminated, each firearm/tool produces a signature of identification (striation/impression) that is unique to that firearm/tool. Through the examination of the individual striations/impressions, the tool mark signature can be positively identified to the firearm/tool that produced it (Freeman, 1978; Hall, 1983; Brundage, 1998; Miller, 2000; Hamby; 2001; Hamby & Brundage, 2007; Hamby, Brundage & Thorpe, 2009; and Fadul, 2011).

Data also revealed no significant differences in the error rate between identifications made by firearm and tool mark examiners with < 10 years of experience (0.008, n = 75) as compared to identifications made by examiners with > 10 years of experience (0.007, n = 90) when examining bullets fired through consecutively manufactured barrels. These results indicate that a trained firearm and tool mark examiner with two years of training, regardless of experience, will correctly identify same gun evidence.

The most significant finding in this study was the low error rate for the examination of unknown bullets and identification to the firearms that fired them when examining bullets fired through consecutively manufactured barrels utilizing individual, unique and repeatable striations. The error rate of the participants was 0.007.

Finally, this research study addressed concerns that were raised by the 2009 National Academy of Sciences Report entitled "Strengthening Forensic Science in the United States: A Path Forward." The National Academy of Sciences Report questioned the repeatability and uniqueness of striations left on fired evidence used to identify same gun evidence as well as questioned the error rate of firearms identification. Based on this research study, firearm and tool mark examiners demonstrated a very low error rate when comparing bullets fired in consecutively manufactured barrels.

Limitations

The researchers noted the following limitations to this study:

- The researchers assumed that the participants followed appropriate AFTE procedures, as listed in the AFTE Procedures Manual, FA-IV-13, Microscopic Comparison (2001).
- Each participant was administered the experimental exercise at their own crime laboratory via mail, and the researchers had no observable control.
- The researchers had to assume that each participant independently completed the experimental exercise with no outside assistance.

- The researchers had no control of the equipment that participants utilized for the experimental exercise.
- The researchers had to assume that the equipment utilized was appropriate, properly maintained and properly functioning.
- The researchers had no control over the training, skill level or experience of the participants.
- The instrument for the experimental exercise was individually administered utilizing the United States Postal Service according to the email response of the participants.
- While the researchers personally mailed the experimental exercise to one participant per crime laboratory, that participant in turn maintained control of the exercise.
- The researchers had no control of the development and maintenance of standards utilized by the participants' laboratories.
- The researchers had no control over participants' laboratory policies regarding inconclusive results, such as whether or not eliminations were allowable based on individual characteristics.
- The study did not follow actual case practice of technical review.

Recommendations for Future Research

Future research is needed in the forensic science community in the area of multiple consecutively manufactured barrels. Research has been conducted on multiple consecutively manufactured barrels (Freeman, 1978; Hall, 1983; Brundage , 1998; Miller, 2000; Hamby; 2001; Hamby and Brundage, 2007; Hamby, Brundage and Thorpe, 2009;

and Fadul, 2011); however, the present research study was the first investigation to utilize multiple participants to examine fired bullets from consecutively manufactured Glock EBIS barrels with the same EBIS pattern in order to determine an error rate. Participants from 125 crime laboratories in 41 states, the District of Columbia, and four international crime laboratories participated in this study; however, additional participants from the remaining crime laboratories and states should be sought out. Future research should include a re-test of the original participants to examine repeatability of the results.

Future research should continue to analyze the repeatability and uniqueness of striations/impressions, as well as examine the reproducibility of the EBIS pattern of other calibers of Glock firearms. Further studies should also continue to incorporate "open set" designs where the participant has no expectation that all unknown tool marks should match one or more of the knowns.

Additional empirical research will continue to support the scientific foundation of forensic firearm and tool mark identification through the evaluation of the uniqueness of striations/impressions and the determination of error rates for the identification of same gun evidence from the additional data. Fundamental research will continue to improve the understanding of the accuracy, reliability and validity of the forensic science discipline of firearm and tool mark identification.

ACKNOWLEDGMENTS

Thank you to the Miami-Dade Police Department Crime Laboratory Firearm and Tool Mark Examiners:

- Chris Barr
- Jorge Bello
- Yamil Garcia
- Angela Garvin
- George Hertel
- Julie Knapp
- John Mancini
- Magen Roberts
- Jill Therriault
- Tim Wilmot

DISSEMINATION OF RESEARCH FINDINGS

South Florida Firearm and Tool Mark Examiners 4th Annual Discipline Meeting, Presentation, February 7, 2013, Fort Pierce, FL.

International Forensic Research Institute (IFRI), 2nd Annual Forensic Science Symposium at Florida International University, Presentation, March 14, 2013, Miami, FL.

ASCLD 40th Annual Symposium, Presentation, May 7, 2013, Durham, NC.

AFTE 2013 44th Annual Training Seminar, Presentation, June 27, 2013, Albuquerque, NM.

REFERENCES

Association of Firearm and Tool Mark Examiners Procedures Manual. Version: July 9, 2001. Association of Firearm and Tool Mark Examiners Member's Page. <http://www.afte.org/MembersArea/promanual.shtml>.

Association of Firearm and Tool Mark Examiners Page. Association of Firearm and Tool Mark Examiners. 16 June 2009 <http://www.afte.org/Membership/membership.htm>.

Brundage, D. (1998). The identification of consecutively rifled gun barrels. *AFTE Journal, 30*(3), 438 – 444.

Butler, J. (2005). Forensic DNA typing, (2nd ed.). Burlington, MA: Elsevier Academic Press.

Carr, J., & Fadul, T. (1997). The Miami barrel. *AFTE Journal*, *29*(2), 232 – 233.

Chin, J., & Sampson, B. (2007, May). *Glock, Inc., Enhanced Bullet Identification System (EBIS) Barrels*. Paper presented at the meeting of the Association of Firearm and Tool Mark Examiners Annual training seminar, San Francisco, CA.

Christensen, L. (2004). Experimental methodology, (9th ed.). Boston, MA: Pearson.

City of New York Police Department. (1996). *Ballistics comparison of polygonal and conventionally rifled Glock, Inc., barrels.* (Communication No. 65/96). New York: Dominic Buccigrossi.

Collins, J. (2009). Stochastics – The real science behind forensic pattern identification. *Crime Lab Report*, retrieved from www.crimelabreport.com.

Creswell, J. (2005). Educational research: Planning, conducting, and evaluating quantitative and qualitative research (2nd ed.). Upper Saddle River, NJ: Pearson.

Fadul, T. (2011). An empirical study to evaluate the repeatability and uniqueness of striations/impressions imparted on consecutively manufactured Glock EBIS gun barrels. *AFTE Journal 43*(1), 37 – 44.

Fadul, T., & Nunez, A. (2003). The Miami barrel saga continues. *AFTE Journal*, *35*(3), 290 – 297.

Fadul, T., & Nunez, A. (2006). Glock, Inc.'s new EBIS barrel: The finale to the Miami barrel. *AFTE Journal*, *38*(2), 96 – 100.

Faigman, D. (2010, August). The Scientific Foundations of Forensic Science. Paper presented at the NIJ Impression and Pattern Evidence Symposium, Clearwater Beach, FL.

Freeman, R. (1978). Consecutively rifled polygon barrels. *AFTE Journal, 10*(2), 40 – 42.

Gall, M., & Borg, W. (1996). Educational research (6th ed.). White Plains, NY: Longman.

Haag, L. (1977), Heckler & Koch polygonal rifling. *AFTE Journal, 9*(2), 45 – 47.

Haag, L., (2003, May). Identifiable Bullets from Glocks in 60 Seconds. Paper Presented at the meeting of the Association of Firearm and Tool Mark Examiners Annual training seminar, Philadelphia, PA.

Haag, L. (2006). RE: Glock's new EBIS barrel by Fadul and Nunez [Letter to the editor]. *AFTE Journal, 38*(3), 189.

Haag, L., Haag, M., Garrett, L., Knell, J., & Patel, M. (2009). Bore lapping compounds and their effect on traditional firearms identification. *AFTE Journal, 41*(3), 232 – 245.

Hall, E. (1983). Bullet markings from consecutively rifled Shilen DGA barrels. *AFTE Journal, 15*(1), 33 – 53.

Hamby, J. (2001). *Forensic firearms examination.* Unpublished doctoral thesis, University of Strathclyde, Glasgow, Scotland, U.K.

Hamby, J., & Brundage, D. (2007, May). *The Identification of Consecutively Rifled 9mm Pistol Barrels*. Paper presented at the meeting of the Association of Firearm and Tool Mark Examiners Annual Training Seminar, San Francisco, CA.

Hamby, J., Brundage, D. & Thorpe, J (2009). The identification of bullets fired from 10 consecutively rifled 9mm Ruger pistol barrels: A research project involving 507 participants from 20 countries. *AFTE Journal (41)*2, 99 – 110.

Hocherman, G., Giverts, P., & Shosani, E. (2003). Identification of polygonal barrel sub-family characteristics. *AFTE Journal, 35*(2), 197 – 200.

Martinez, C. (2008, May). *Glock's Signature Barrel – Durability of the EBIS Markings.* Paper presented at the meeting of the Association of Firearm and Tool Mark Examiners Annual Training Seminar, Honolulu, HI.

Miller, J. (2000). An examination of two consecutively rifled barrels. *AFTE Journal, 32*(3), 259 – 270.

Mnookin, J. (2010, August). Show me the Data. Paper presented at the NIJ Impression and Pattern Evidence Symposium, Clearwater Beach, FL.

Moran, B., & Murdock, J. (2002). Joseph Ramirez vs. State of Florida – Supreme Court Decision December 20, 2001. *AFTE Journal (34)* 2, 215 – 226.

National Academy of Sciences. Strengthening Forensic Science in the United States: A Path Forward. Washington, D.C.: The National Academies Press, 2009.

Nichols, R. (2007). Defending the scientific foundations of the firearms and tool mark identification discipline: Responding to recent challenges [Abstract]. *Journal of Forensic Sciences, 52*(3), 586 – 594.

Northcutt, A. (2010), "Glock barrel modification: a feasible method to facilitate the identification of fired bullets. *AFTE Journal, 42*(2), 108 – 119.

Peterson, J., & Markham, P. (1995). Crime Laboratory Proficiency Testing Results, 1978-1991, II: Resolving Questions of Common Origin. *Journal of Forensic Sciences, 40*(6), 1009-1029.

Saks, M. (1993). The Unknown Validity of forensic Science (draft). College of law, University of Iowa. 1 – 21.

Saks, M. (1994). Implications of the Daubert Test for Forensic Identification Science. *Shepard's Expert & Scientific Evidence,* (1)3, 427 – 434.

Saks, M., & Koehler, J. (1991). What DNA Fingerprinting Can Teach the Law About the Rest of Forensic Science. *Cardozo Law Review, (13)*2-3, 361 – 372.

Saks, M., & Koehler, J. (2008). The Individualization Fallacy in Forensic Science Evidence. *Vanderbilt Law Review, (61)*1, 199 – 219.

Valdez, S. (1997), Bullet identification from H & K USP polygonal barrels. *AFTE Journal,* 29(3), 307 – 309.

Appendix A:

Survey Instrument (Questionnaire/Answer Sheet)

Miami-Dade Police Department
Crime Laboratory
9105 NW 25th Street, Miami, Florida 33172
(305) 471-2050

Firearm & Toolmark Unit

Answer Sheet: Consecutively Rifled **EBIS-2** Test Set Test Number: __________

Name: **ANONYMOUS** Male or Female? (Please circle one) Date:_________________

Years Experience:______________Years Training:____________Type of Training:__________________________

Brand & Model of Microscope:__________________________Type of Lighting:____________________________

QCMS Trained? Yes No Did you use Pattern Matching, QCMS or Both for this test?__________________

Is your Laboratory ASCLD/Lab Accredited? Yes No Other Accreditatation? _____________________

AFTE Certified? Yes No ABC Certified? Yes No Other Certification? ___________________

Have you ever encountered the Miami or EBIS Barrel in your case work? Yes No If yes, How many times?_____

Did you participate in the First Miami Barrel / EBIS Study? Yes No

Please microscopically compare the known test shots from each of the 8 barrels with the 10 questioned bullets submitted. Indicate your conclusion(s) by circling the appropriate known test fired set number designator on the same line as the alpha unknown bullet. You also have the option of Inconclusive and Elimination. This test does not have to be done all at one time, but sufficient time to adequately examine this material is necessary. Although the bullets have been scribed on the nose, you may elect to confirm the 'identifier' on the nose and re-scribe it on the base of the bullet.

Unknowns **Knowns** (Barrels 1 through 8)

A. 1…....2…….3..…..4…….5…….6..…..7…....8…....Inconclusive……Elimination

B. 1…....2…….3..…..4…….5…….6..…..7…....8…....Inconclusive……Elimination

C. 1…....2…….3..…..4…….5…….6..…..7…....8…....Inconclusive……Elimination

D. 1…....2…….3..…..4…….5…….6..…..7…....8…....Inconclusive……Elimination

E. 1…....2…….3..…..4…….5…….6..…..7…....8…....Inconclusive……Elimination

F. 1…....2…….3..…..4…….5…….6..…..7…....8…....Inconclusive……Elimination

G. 1…....2…….3..…..4…….5…….6..…..7…....8…....Inconclusive……Elimination

H. 1…....2…….3..…..4…….5…….6..…..7…....8…....Inconclusive……Elimination

I. 1…....2…….3..…..4…….5…….6..…..7…....8…....Inconclusive……Elimination

J. 1…....2…….3..…..4…….5…….6..…..7…....8…....Inconclusive……Elimination

Other Results/Comments: ____

Adapted from the Indianapolis-Marion County Forensic Services Agency with the permission of Dr. James E. Hamby

The author(s) below used Federal funds provided by the U.S. Department of Defense and prepared the following report:

Document Title:	A Study of False-Positive and False-Negative Error Rates in Cartridge Case Comparisons
Author(s):	David P. Baldwin, Stanley J. Bajic, Max Morris, and Daniel Zamzow
Document No.:	NCJ 249874
Date Received at NCJRS:	May 2016
Contract Number:	DE-AC02-07CH11358

A Study of False-Positive and False-Negative Error Rates in Cartridge Case Comparisons

David P. Baldwin, Stanley J. Bajic, Max Morris, and Daniel Zamzow

Ames Laboratory, USDOE

Technical Report # IS-5207

April 7, 2014

Office of the CHIEF SCIENTIST
DEFENSE FORENSIC SCIENCE CENTER

Sponsors and Disclaimer

This work was supported by Defense Biometrics and Forensics Office, Assistant Secretary of Defense (Research and Engineering), through the U.S. Department of Energy Contract No. DE-AC02-07CH11358. The views and conclusions contained in this document are those of the authors and should not be interpreted as representing the official policies, either expressed or implied, of the Defense Biometrics and Forensics Office, Defense Forensic Science Center, or the U.S. Government. The U.S. Government is authorized to reproduce and distribute reprints for Government purposes notwithstanding any copyright notation hereon.

Acknowledgements

This work was supported by Defense Biometrics and Forensics Office through the U.S. Department of Energy under Contract No. DE-AC02-07CH11358. The study would have been impossible without the willing participation of the hundreds of examiners and the permission of their agencies to participate voluntarily. The study was managed by the Defense Forensics and Biometrics Agency and the Defense Forensic Science Center, Office of the Chief Scientist. Henry Maynard, Garold Warner and other members of the OCS staff located at the Defense Forensic Science Center shaped and guided the project. We give special thanks to Dr. Michael "Jeff" Salyards, Executive Director, DFSC, Rick Tontarski, DFBA /DFSC Chief Scientist, and the staff of the firearms section at USACIL, a division of the DFSC. The Forensic Initiative at WVU provided the firearms, ammunition, and support in firing and collecting samples for the study including Dr. Keith Morris, Mike Bell, and several student staff members. The Forensic Research Committee of ASCLD provided input and guidance as well as valuable recruitment support for both the pilot study and the full study and we want to particularly thank Jay Henry for his efforts in communications through ASCLD membership. AFTE provided the largest group of participants through the support and communications provided by Jay Stuart. The Story County, Iowa Sheriff's Office made the pilot study possible and we want to particularly thank Sheriff Fitzgerald, Sergeant Backous, and Detective Rhoads. SWGGUN Chair Andy Smith provided valuable guidance in designing the study and aided in identifying the proper selection of ammunition. The staff and management of the Firearms & Toolmarks Unit at the FBI Laboratory provided additional input on project design. Finally we thank the staff of the Ames Laboratory Shipping and Receiving Department for coordinating and handling shipment of all of the materials for the study, in particular Vicki Sieve.

Abstract: This report provides the details for a study designed to measure examiner (not laboratory) error rates for false identifications and false eliminations when comparing an unknown to a collection of three known cartridge cases. Volunteer active examiners with Association of Firearm and Toolmark Examiners (AFTE) membership or working in laboratories that participate in ASCLD were provided with 15 sets of 3 known + 1 unknown cartridge cases fired from a collection of 25 new Ruger SR9 handguns. The ammunition was all Remington 9-mm Luger (manufacturer designation L9MM3) and sets were made up of cartridge cases fired within 100 cartridges of each other for each gun. During the design phase of the experiment, examiners had expressed a concern that known samples should not be separated by a large number of fired cartridges. However, studies published on this effect indicate that several thousands of cartridges could be fired by the same firearm without making the identifying characteristics change enough to prevent identification. [1] Examiners were provided with a background survey, an answer sheet allowing for the AFTE range of conclusions, and return shipping materials. They were also asked to assess how many of the 3 knowns were suitable for comparison, providing a measured rate of how often each firearm used in the study produces useable, quality marks. The participating examiners were provided with known positives and known negatives from independent groups of samples, providing independent measurements of a false-positive rate and independent measurements of a false-negative rate, allowing the study to measure both rates and uncertainties in those rates.

Responses were received from 218 participating examiners. The rate of false negatives (estimated as 0.367% from comparisons known to be from the same firearm but reported as eliminations) was quite low with the error distributed across examiners of various backgrounds (state, federal, local, private, etc. as determined from self-reported survey information). The overall rate of false positives (estimated as 1.01% from comparisons known to be from different firearms but reported as identifications) was significantly higher. However, most of the errors were reported by a small number of examiners; that is, individual examiners have varying error rates. For most examiners this is quite low while for some it is relatively high. Hence the overall rate is best interpreted as an average of widely varying individual rates. Inconclusive results were not recorded as errors. Rates of poor quality mark production for these handguns varied across the 25 sample handguns. Those rates were 2.3 (±1.4) %.

False-positive and false-negative error rates for individual examiner performance on comparisons were measured. The rates are not uniform across the sample population with a few examiners providing most of the false-positive responses. False-negative rates are low and comparable to or lower than the rate of production of poor quality marks by the firearms used in this study. Laboratory error rates may be significantly lower than these individual rates if quality assurance procedures are applied that can effectively manage to reduce or eliminate the propagation of false positives reported by individuals.

Introduction:

This study was designed to provide a better understanding of the error rates associated with the forensic comparison of fired cartridge cases. Several previous studies have been carried out to examine this and related issues of individualization and durability of marks [1-5], but the design of these previous studies, whether intended to measure error rates or not, did not include truly independent sample sets that would allow the unbiased determination of false-positive or false-negative error rates from the data in those studies.

The Admissibility Resource Kit (ARK) developed and published by the Scientific Working Group for Firearms and Toolmarks (SWGGUN) maintains an extensive bibliography of literature relevant to the reliability, repeatability, and validation of forensic examinations of cartridge cases. Currently, this bibliography can be found at:

http://www.swggun.org/swg/index.php?option=com_content&view=article&id=5:testability-of-the-scientific-principle&catid=9:ark#CC

Please note that there are tentative plans to transfer the maintenance of the ARK to AFTE in the event that SWGGUN is replaced with another standards organization.

The study was designed with sample sets for comparison that are as independent as economically feasible given the cost of firearms and ammunition. We set out to measure both false-positive rates and false-negative rates, so the participants were presented with 15 independent comparison sets that we knew either came from the same source or didn't. No source firearm was repeated within any participant's test packet except within a set that was from the same source. This was deemed important since there was anecdotal feedback during discussions of design that examiners remember patterns even outside a posed comparison, and it was reported that this might affect the responses or the perception of the experimental design. Initially we had considered designing each set as a 1 to 1 comparison. We obtained feedback that some labs do not allow this type of analysis so we decided instead to provide examiners with 3 "known" cartridge cases to compare to 1 "questioned" case. This was meant to mimic the situation where an examiner would compare a questioned case to repeated firings from a firearm in evidence. The provision of 3 knowns in this study allowed us to address an issue of determining how often each firearm in the study would produce a fired cartridge case with insufficient markings for a comparison. In each set, the participants were asked first to determine how many of the 3 knowns had marks that were suitable for comparison. One valuable outcome of this study is an extensive measurement of this poor mark production rate for the 25 SR9 handguns used in this study. It is important to remember that these poor marking rates might well be different with different models of firearm, firearms of significantly different age or condition, and different makes and manufacture of ammunition. However, our data provides a reliable measurement of this phenomenon for the samples used in this study. The measurement of this poor marking rate allows us to understand the potential effects on the measured examiner error rates. By measuring this rate of poor mark reproduction, we also avoid a problematic practice of prescreening the quality of samples provided to the participants. In some previous studies, a qualified examiner would screen all samples to make sure that only well-marked samples were included. This introduces two problems. One is that the study risks the criticism that only "easy" samples are included, which is not necessarily reflective of real casework samples. The second is that even if the screener is very qualified (and can afford or has the fortitude to examine the 20,000 cases used in this study) the judgment of quality marks is a subjective judgment best left to each examiner. By instead allowing all of the examiners to report a rate of poor production for their known samples we gain an insight into the effect without having to accept the cost and ambiguity of a reanalysis for every sample.

A very important aspect of this work that needs to be clearly understood is that the study specifically asked participants not to use their laboratory or agency peer review process. There are two aspects to consider in this decision. First, the errors that are identified during the study and the estimated rates of error are for individual examiner performance and not directly related to rates of error for reported analyses from any laboratory or agency. The variability in how, how often, and how effectively the quality assurance programs in the dozens of agencies in which the participants practice might have detected and prevented reporting errors is so wide and undefined that we deemed a study that included these QA measures as being less meaningful than a study of how well trained individuals perform in a relatively uniform set of circumstances. After reviewing the results of this study, we would hope that QA managers would institute appropriate controls to manage the measured performance error rates, rather than taking a snapshot of the systemic performance of existing disparate systems without understanding what errors are being managed by those systems.

The firearm chosen for this study is the Ruger SR9 semiautomatic 9-mm handgun. This is a relatively new model of handgun, but it was chosen for several reasons. Sturm, Ruger & Company is a popular firearms brand with a reasonably positive reputation for service and quality. Although not the most popular source for semiautomatic handguns that might be chosen as a service sidearm compared to industry giants like Glock, Sig Sauer, Beretta, and S&W, Ruger has an apparent sizeable share of the market for home defense weapons. In part this is due to the lower cost of their handguns compared to similarly featured products from these other manufacturers. At the time we began this study, the retail prices for the Ruger SR9 pistol were as much as 25 to 30% lower than a comparable 9-mm Glock pistol, making a study using the Ruger SR9 more affordable. For this same reason, Ruger products may be more likely to be obtained and available for illegal activities. Currently, the Ruger P95 is a more mature product line that exists in larger numbers than the Ruger SR9. However, Ruger was discontinuing production of their P-series firearms, and have since halted production of the P95 (in October 2013). The Ruger SR series is fully replacing the P series as their full size semiautomatic handgun line. We would expect the SR series to eventually rival if not surpass numbers of P-series firearms in circulation. An additional reason for choosing the Ruger SR9 was the need for reliable performance for production of the quantity of samples needed for this study. The striker-fired design, similar to that employed by the popular Glock line, has proven to be quite reliable compared to other hammer-based designs and uses fewer moving parts. While there were several anecdotal suggestions that cartridge cases from firearms like the Glock might have been "too easy" or the Bryco/Jennings/Jimenez style handguns might be "too hard" or have too high a rate of poorly reproduced marks, the Ruger SR9 had no such preconceived biases for or against its use. This study incorporates a design to measure the reproduction rate for useable comparison marks, and the relative ease of comparison is an issue that will always need to be addressed to generalize this type of study to more than one model of firearm. We chose to base the available responses on the AFTE Range of Conclusions, given the preponderance of AFTE members as participants in the study. Therefore, we interpret the meaning of inconclusive according to the AFTE guidelines that establish this range of conclusions (http://afte.org/AssociationInfo/comm & info/roc.htm – also reproduced below). If the examiner does not find sufficient matching detail to uniquely identify a common source for the known and questioned samples, and there are no class characteristics such as caliber that would preclude the cases as having been fired from the same-source firearm, a finding of inconclusive is an appropriate answer (and not counted as an error or as a non-answer in this study). The underlying rationale for this finding of inconclusive is that the examiner is unable to locate sufficient corresponding individual characteristics to either include or exclude an exhibit as having been fired in a particular firearm and the possible reasons are numerous as to why insufficient marks exist. As is determined in this study, there are also a significant number of times that the firearm fails to make clear and reproducible marks (which very well might have happened for a questioned case). On the other hand, we received many conflicting comments from study participants regarding the use of this range of conclusions, indicating systemic differences in how closely these guidelines are followed or how the guidelines are interpreted. Some participants indicated that they would not participate if the full range of conclusions, including different classifications of conclusions for inconclusive findings, were not included. Other participants indicated that their agencies did not allow them to distinguish between the different classifications for inconclusive results so they chose not to differentiate between any of the three options within the AFTE guidelines in their answers. Some indicated that the design of our study with all cartridges fired from the same model of firearm using the same type of ammunition would prohibit the use of a finding of elimination, while others used a mixture of inconclusive and elimination or did not use inconclusive at all to indicate a finding other than identification. In this report we present our findings on these variations and suggest that this is an area for further study and discussion within the community.

One of the reasons for undertaking this work was to address an issue identified in the 2009 National academies report: “Strengthening Forensic Science in the United States: A Path Forward.”[6] In the summary assessment section on toolmarks and firearms identification in that report the authors stated that: “Sufficient studies have not been done to understand the reliability and repeatability of the methods.” This work addresses that issue through an assessment of the reliability of the comparison of fired cartridge casings.

AFTE Range of Conclusions, reproduced from http://afte.org/AssociationInfo/comm & info/roc.htm.

The examiner is encouraged to report the objective observations that support the findings of toolmark examinations. The examiner should be conservative when reporting the significance of these observations.
The following represents a spectrum of statements:

1) IDENTIFICATION

Agreement of a combination of individual characteristics and all discernible class characteristics where the extent of agreement exceeds that which can occur in the comparison of toolmarks made by different tools and is consistent with the agreement demonstrated by toolmarks known to have been produced by the same tool.

2) INCONCLUSIVE

A. Some agreement of individual characteristics and all discernible class characteristics, but insufficient for an identification.

B. Agreement of all discernible class characteristics without agreement or disagreement of individual characteristics due to an absence, insufficiency, or lack of reproducibility.

C. Agreement of all discernible class characteristics and disagreement of individual characteristics, but insufficient for an elimination.

3) ELIMINATION

Significant disagreement of discernible class characteristics and/or individual characteristics.

4) UNSUITABLE

Unsuitable for examination.

Finally, the study was limited to participants who are self-reporting as active forensic firearm examiners who are either AFTE members or work in laboratories that participate in ASCLD. Although it might be desirable to understand how non-practicing or untrained participants might perform under the same circumstances as trained examiners, there are important statistical reasons for not including trainees. The expected rates of error are low enough that dividing our participant pool into subgroups that are trained and not trained would add cost to the study without adding enough participants to allow a precise measurement of error rates for this group of trainees. It was deemed more important to measure the error rates for trained practicing examiners accurately and precisely than to measure the effect of another variable with much less precision and accuracy.

Experimental:

Pilot study:

Before embarking on soliciting participants and assembling test materials for the full study, we conducted a pilot test to solicit input from volunteers from four laboratories recruited through the ASCLD Forensic Research Committee (FRC). Ammunition and firearms purchases for this study were delayed for several months by contractual arrangements and ongoing market supply issues at the time. As a result, test samples were obtained through the voluntary cooperation of a local police agency. The test sets (240 fired cases) were collected by the Story County, Iowa Sheriff's Department using five of their service sidearms (.40 S&W caliber, Sig Sauer P229), obtained during a training session for that department. Four pilot test cartridge case packets were assembled and sent to volunteers at the four volunteer agencies. We requested feedback on the study design, survey forms, informed consent forms, and any other information they wanted to provide. Several improvements to the study materials were made as a result of this pilot study. One significant limitation of the pilot study was the limited number of firearms available to assemble the test sets, resulting in many more repeated source samples than were used in the full study.

Test firing and Selection of Ammunition:

Among those consulted during the design phase of this study was Andrew Smith, SWGGUN Chair and a firearms examiner with the San Francisco Police Department. He encouraged us to test fire several types of ammunition with the model of handgun to be used in the study. We had obtained three test fired samples of six commonly available 9-mm cartridges from Blazer, Winchester (two types), Federal, Remington, and PMC and sent them to Mr. Smith for evaluation. All cartridges produced useable detail.

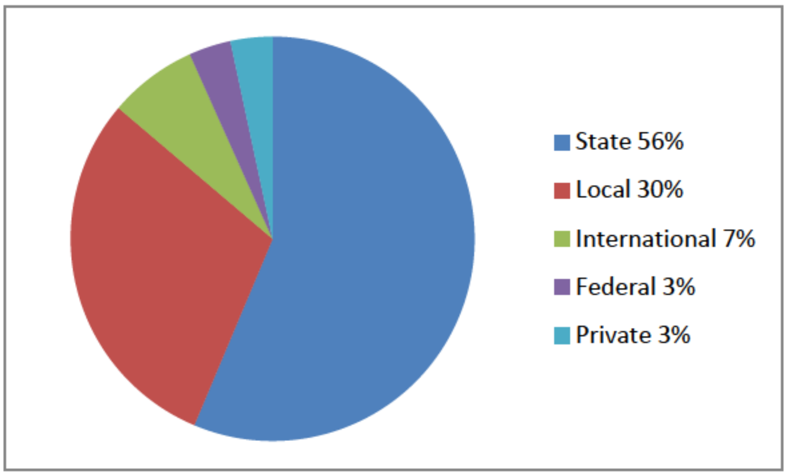


Figure 1: Distribution of the 284 respondents enrolled in the study.

Solicitation of Participants:

Due to the use of human subjects in this study, the experimental program was subject to review and approval by Institutional Review Boards at Ames Laboratory (using the Institutional Review Board of our contracting agency, Iowa State University) and by the Department of Defense. The study was designed to protect the participants from risk to their professional standing and reputation by making all results strictly anonymous. One group of researchers prepared materials and scored the responses and a separate group solicited participants and sent and received sealed unlabeled study material packets. No researcher has information connecting identities with scored results.

Invitations to participate in this study were broadcast by email and newsletter publications through ASCLD and AFTE. The invitation letter and the informed consent form for this study are included in Appendix A. The study accepted registrations to participate during the months of August and September of 2013. We received consent forms and issued test materials to 284 participants and received completed results from 218 respondents by the end of December, 2013. Answers and identities, including significant identifying information like home organization, have been completely separated. However, from our respondent pool we can determine that 56% of those enrolled were employed by domestic state-level agencies, 30% were employed by domestic local agencies, 7% were employed by international (non-U.S.) agencies, 3% were employed by domestic federal agencies, and 3% were self-employed or worked for private forensic agencies.

Figure 2: A Ruger© SR9© handgun with two 17-round magazines. This is the same model as the firearms used in this study.

Sample Preparation and Collection:

Given restrictions on acquisition and property management of firearms by the Ames Laboratory, our collaborators at the Forensic Science Initiative at West Virginia University (WVU) purchased twenty-five (25) new Ruger SR9, 9mm Luger centerfire pistols. The serial numbers and assigned identifiers (explained below) used for each handgun in this study are listed in the Table in Appendix B. WVU also purchased 25,000 cartridges of Remington UMC Pistol and Revolver Cartridges. The cartridges used were 115 grain full metal jacket (FMJ) bullets. The ammunition came in boxes of 50 cartridges with each box containing a 50-holed plastic bullet carrier to keep the individual cartridges from "bouncing around" during shipping and handling, potentially acquiring additional marks. The ammunition came from two lots from the manufacturer: 136 boxes from Lot# 3503 ODEL 2195989 and 264 boxes from Lot# 350 ODFL 32206416. Prior to the sample collection campaign for this study, staff from WVU fired 200 cartridges of ammunition using each handgun and collected the fired cases for use in an unrelated study. There is anecdotal evidence to suggest that there is some rapid early change in marks made in the first several firings with a new handgun. The cartridges were collected by WVU to investigate this effect in a separate study. For the same reason, these fired cartridge cases were not used in this error-rate study.

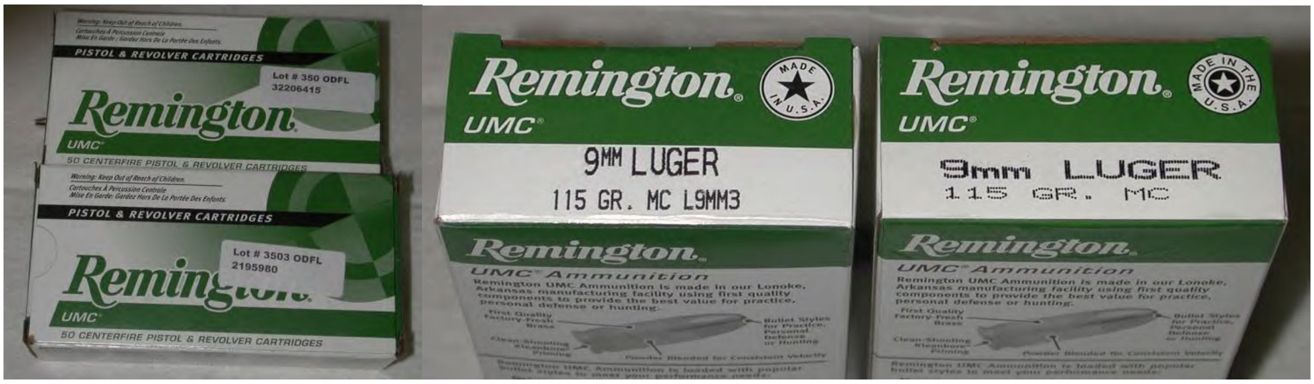


Figure 3. Boxes containing ammunition used in this study.

The ammunition was divided up into batches of 16 boxes, which correspond to 800 cartridges per batch. Each batch was allotted to one of the 25 Ruger SR9 handguns. Except for handgun "B1," each handgun was used to fire all 800 cartridges of ammunition from the same Lot. Handgun "B1" was used to fire 400 cartridges of ammunition from each of the two lots.

The sample sets for the study were made up of 15 comparisons of 3 knowns to 1 questioned cartridge case. For all participants, 5 of the sets were from known same-source firearms, and 10 of the sets were from known different-source firearms. We instructed all participants to refrain from sharing or discussing the contents or results of their sample sets and answers to minimize the risk of revealing this design. We did not share these details to anyone outside the group assembling the test sets and the local project managers. In order to provide 5 independent sets of samples to measure a rate of false-negative conclusions, which could then be used to determine not only a rate but an uncertainty in that rate, the 25 handguns were broken into 5 groups of 5 handguns (A1-A5, B1-B5, C1-C5, D1-D5, E1-E5). Each participant was randomly assigned to one of the groups A through E. If assigned to group A, that examiner would receive 5 sets where cartridges from handgun A1 would be used for all of the cartridge cases in one set through A5 making up the fifth set of all same-source cartridge cases. These designations did not appear on any markings provided to the participants and the order within the 15 sets within a test packet was randomized. The actual identity of the sources and the assembly of the test kits and was tracked by an individual on the project team without knowledge of the identity of the participants, and another individual checking the accuracy of his work (also not privy to the identity of the participants). The remaining 10 sets for each examiner were assembled from different-source firearms from different groups. Those in the A group were also provided 5 sets with firearms from the B group producing the knowns and D firearms producing the questioned cartridge cases, and 5 sets with firearms from C producing the knowns and E firearms producing the questioned cartridge cases. In order to further randomize the sets between examiners in a given group, the pair of firearms in the different-source comparison sets was incremented for each examiner so that only every sixth participant in a group would receive the same comparison group. So the A group of examiners saw A(k) v. A(q), B(k) v. D(q), and C(k) v. E(q) (likewise the remaining groups were assigned as B group: B-B/C-E, D-A; C group: C-C/ D-A, E-B; D group: D-D/ E-B, A-C; E group: E-E/ A-C, B-D). The net effect of this design is that every participant received cartridges from each of the 25 firearms as either a known or questioned sample (or both for known same-source sets), each known same-source group was independent of the other four groups, and no firearm was used more than once for a participant unless it was part of a known same-source set.

Table I: Experimental Set Design

Group	A	B	C	D	E
Same-source comparisons	A1-A1	B1-B1	C1-C1	D1-D1	E1-E1
	A2-A2	B2-B2	C2-C2	D2-D2	E2-E2
	A3-A3	B3-B3	C3-C3	D3-D3	E3-E3
	A4-A4	B4-B4	C4-C4	D4-D4	E4-E4
	A5-A5	B5-B5	C5-C5	D5-D5	E5-E5
Different-source comparisons	B v D: 1v2, 2v3, 3v4, 4v5, 5v1 and other skip permutations	C v E	D v A	E v B	A v C
	C v E	D v A	E v B	A v C	B v D

False-negative error rates for each of the 5 same-source groups (A through E in Table I) provide 5 independent measurements of this rate and allow the measurement of uncertainty in the false-negative rate. False-positive error rates for each of the 5 different-source groups (B v. D, C v. E, D v. A, E v. B, and A v. C in Table I) provide 5 different independent measurements of the false-positive rate and allow the measurement of uncertainty in the false-positive rate.

Cartridge cases were collected in groupings of 100 with 800 total cartridge cases collected per firearm. From a batch for a particular firearm, 100 cartridges of ammunition were loaded into six magazines; five magazines loaded with 17 cartridges and one magazine loaded with 15 cartridges. Loading of the magazines took place in the observation area of the firing range. Once loaded, the six magazines were taken into the firing range and used. The handgun was discharge through a homemade brass catcher (Figure 4) with the cartridge cases collected by the catcher after ejection from the firearm. After firing all 100 cartridges, the cartridge cases were removed from the catcher, placed in a sealable plastic bag, and then taken out of the firing range and into another section of the observation area. The cartridge cases were then placed in 100 round plastic ammunition carrier boxes to prevent any additional marks being made on the cases and to facilitate sorting and transporting.

Figure 4. Brass catcher used to collect samples for this study.

The ammunition was fired and collected sequentially (i.e., 1-100, 101-200… 701-800). Only cartridge cases caught by the catcher were collected. If a cartridge case somehow fell out of the catcher, it was immediately discarded. Misfires were cleared from the firearm and disposed of without attempting to reuse the round. The 100-round ammunition carrier boxes were labeled with the firearm letter designation and the segment of cartridges that were collected (e.g., "Gun C2, 501-600").

The firearms were cleaned prior to being used at the range and cleaned at the range after 400 cartridges were fired through the handgun. Cleaning consisted of disassembling the handgun, letting it cool; brushing the breach face with a plastic bristled brush; and running a cotton bore cleaning swab, wetted with Hoppe's No. 9, through the barrel. The firearm was then reassembled and either used to fire the remaining ammunition for that handgun or placed in its carrying case for storage.

Each collected cartridge case was labeled with a unique alphanumeric identifier. The alphanumeric identifier had the format of either "Kmfrcxxxyyy" or "Qmfrcxxxyyy", where "x" is alphabetic letter and "y" is an integer. "K" and "Q" indicate that the labeled cartridge case would be used as a "known" or "questioned" casing in the generated sample sets. This was included so that if a participant were to mix up the K's and Q's within a set, they could still be identified. The 20,000 unique alphanumeric identifiers were randomly generated using a free online random-code generator (http://www.mindflow.com.au/Custom_CMS/index.php/Online-Tools/Random-Code-Generator-Free-Tool.html). The output from the random-code-generator was a csv format file that could be imported into a label making program.

Labels for the casings were generated using a Brother P-touch PT-2430PC Label Maker on ¼" white tape with black print (TZe-211) that was laminated. The printed labels were then manually affixed to each individual cartridge case. Figure 5 depicts examples of labeled cartridge cases. Although the labels as affixed have in some instances covered the ejector marks on the cartridge cases, the participants could have removed this label if necessary to make a definitive examination. There is some possibility that participants may have felt restrained from doing so for some reason and may have performed the examinations without information they normally would have used. We received only one comment to this effect. The labels were placed in this way due to the difficulty, time, and expense involved in placing the labels in a less obtrusive location, such as inside the cartridge cases. Had the study been significantly smaller (with fewer samples to label, to assemble into sets, and check for accuracy in assembly), we would have considered different placement options.

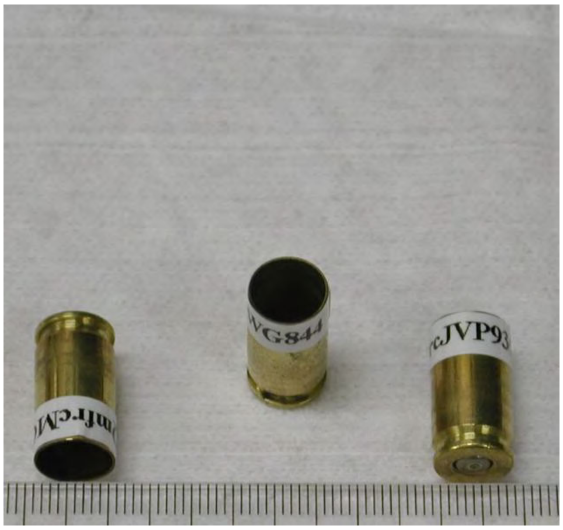


Figure 5. Labelled cartridge cases.

The sequence of the sets in each group was determined by a random number/list generator so that the order of same-source and different-source sets was random for each participant. The generator can be found at (http://www.random.org/lists/).

The response form is shown in Appendix C. Each form was marked with an identifier to associate it with the materials provided so that responses could be compared to known sources. The form includes background demographic information, the full range of AFTE-approved responses, and an evaluation of how many of the knowns in each set were suitable for comparison.

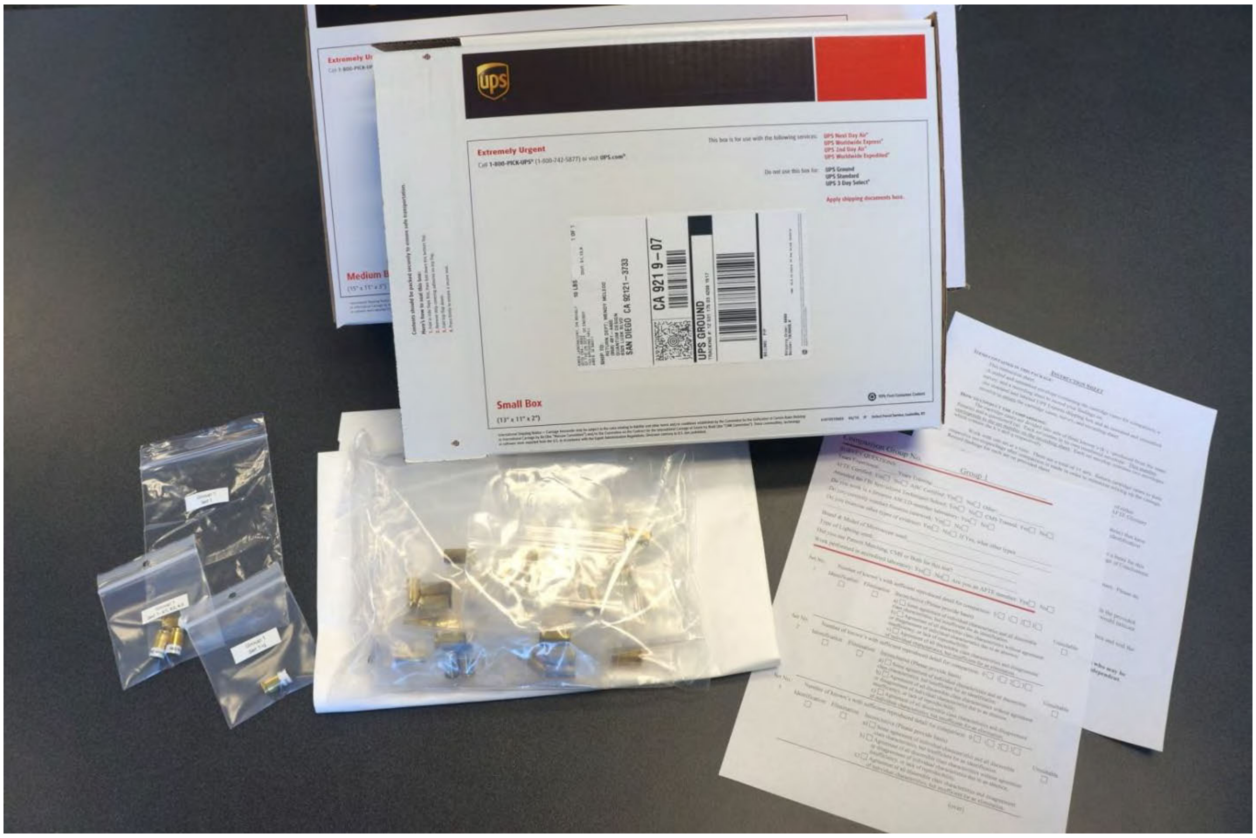


Figure 6. A sample package with supplied samples, return packaging, and forms.

Sample packages were assembled for each participant and included an instruction sheet (see Appendix D), a survey and answer sheet (see Appendix C), a return prepaid shipping box, an unmarked Tyvek envelope for returning the answers and samples, and plastic zip seal bags containing the sample sets marked with a set number (with knowns and questioned cartridge cases also in separate zip bags). The answer sheet and samples were marked with an identifier so that they could be recorded and scored upon return. The shipping box had a separate unrelated identifier that was unassociated with the sample set but connected to the respondent's identity. When materials were returned, that identifier was used to note that a participant was finished and should be removed from reminder mailing lists. The returned samples and answer sheets, sealed in the unmarked Tyvek envelope, were transferred to the scoring group – at which point the identity of the respondent would be separated from the results of the examination.

Several of the participants in the study were working outside the United States. In order to send sample sets to these participants and comply with U.S. export control and arms control regulations, the samples needed to be altered in a way that would prevent reloading of the fired cartridge cases, while still permitting examination. For these sample sets, each sample was cut with an approximately 3-mm long cut along the case wall using a handheld rotary tool with a cutting wheel. Examples of these altered cartridge cases are shown in Figure 7.

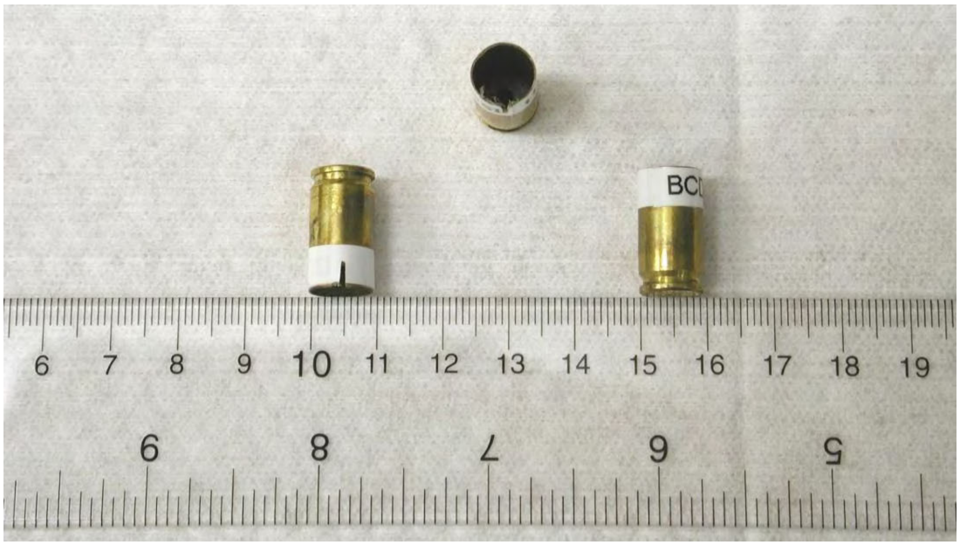


Figure 7. Sample cartridge cases intentionally altered for shipment to international participants.

Results and Discussion:

Of the 1090 true same-source comparisons made (i.e. where the three knowns and the single questioned samples actually were produced by the same firearm), only four comparisons were labeled elimination and eleven were judged inconclusive. The proportion of false elimination calls, relative to all responsive answers (elimination, identification, and inconclusive) was 4/1090 = 0.003670, or 0.3670%. The Clopper-Pearson exact 95% confidence interval [7] for this proportion is (0.001001, 0.009369), or (0.1001%, 0.9369%). Two of the four incorrect elimination calls were made by the same examiner; 215 of the 218 examiners did not make any false elimination calls. The proportion of false elimination and inconclusive calls together (not all errors, but also not identifications), relative to all responsive calls, was 15/1090 = 0.01376, or 1.376%; the corresponding 95% confidence interval is (0.007722, 0.02260), or (0.7722%, 2.260%). For those unfamiliar with statistical methods for confidence interval calculation, please note that Clopper-Pearson is a widely used method of calculating the confidence interval for a true value of a mean from sample data. Using a different method for confidence interval calculation would not affect the reported mean value of the error rate and would result in only minor changes in the size of the range of the confidence interval. Comparing these differences in confidence interval calculations is irrelevant to the discussion or the implications of this report.

One relevant number for comparison to the false elimination probability is the proportion of known samples that were judged to be poorly marked, and so not used by the examiners. Of the 3234 comparisons for which the number of useable knowns was reported (some examiners either did not provide this information or provided it for only some of the samples), all three specimens were used in 3018 cases, two were used in 207 cases, and only one was used in nine cases. Hence the "raw" proportion of known specimens judged by the examiners to be inappropriate for inclusion in the comparison was 225/9702 = 0.02319, or 2.319%, with a corresponding 95% confidence interval of (0.02174, 0.02827), or (2.174% to 2.827%). This is substantially greater than the estimated proportion of false eliminations, and raises the possibility that at least some of these incorrect eliminations may have been the result of poorly marked questioned samples. In addition, compared to the rate of false eliminations and inconclusive results for same-source samples, above, this rate of poorly marked cases may very well explain most if not all of these calls. At best this rate of poor marking makes determination of a false-negative error rate difficult since it is at least comparable in magnitude and indistinguishable in result.

Of the 2180 true different-source comparisons possible (i.e. where the three knowns and the single questioned sample actually were not produced by the same firearm), 22 comparisons were labeled identifications. The remaining correct responses included 735 reported as inconclusive and 1421 as eliminations. The proportion of false identification calls from among all responsive answers was 22/2178 = 0.01010, or 1.010% (two comparisons were not reported or left blank – reducing the total responses from 2180 to 2178). All but two of the 22 false identification calls were made by five of the 218 examiners, strongly suggesting that this error probability is not consistent across examiners (or in effect, that each examiner has his or her own false identification probability, and that these probabilities vary substantially).

As a result, appropriate statistical modeling for the proportion of false identifications is more intricate because it cannot be assumed that the probability of such a call is uniform across the participating examiners (and so the dataset). The beta-binomial model was used as a basis for estimation here; it is a mixture of binomial distributions (appropriate for each individual examiner) in which the individual binomial probabilities follow a beta distribution (representing the variation among examiners). Because the individual error probabilities can only be estimated with substantial uncertainty from the small sample of comparisons made by each examiner, the inference is less precise than the Clopper-Pearson methodology used for false eliminations. The maximum likelihood estimator for the false identification probability, averaged across examiners, is 0.00939 (0.939%), with a likelihood-based 95% confidence interval of (0.00360, 0.02261) or (0.360%, 2.261%).

Examiners were also highly heterogeneous in the number of true different-source cartridge cases they identified as inconclusive. Of 218 examiners, 96 (the largest subcategory) labeled none of the comparisons as inconclusive (instead using elimination to denote a determination that found insufficient matching detail for an identification), while 45 (the second largest subcategory) labeled all 10 of the comparisons as inconclusive. The remaining 77 examiners were fairly evenly spread between these two extremes, with somewhat more of them reporting relatively fewer inconclusive calls. There are mild inverse correlations between the number of inconclusive/nonresponse calls made with the known different-source cases, and the reported number of years of training (correlation = -0.1393) and number of years of experience (correlation = -0.1034); that is, there is a weak tendency for examiners with more training or experience to make fewer inconclusive calls. The distributions of number of inconclusive calls were not substantially different between examiners who identified themselves as AFTE members (183) and those that said they were not (7), examiners who identified themselves as working in ASCLD participating agencies (183) and those that said they were not (29), and examiners that were identified as AFTE certified (48) and those that were not (164). Not every respondent answered every survey question.

Table II. Data Collection Information

Enrolled participants	284
Responding participants	218
Known same-source responses (5x218)	1090
Known different-source responses (10x218 – 2 blanks)	2178
Suitability of knowns responses (3x15x218 – 108 blanks)	9702
False positives	22
False negatives	4
Inconclusive responses for known same-source comparisons	11

Table III. Rates and Confidence Intervals

	Rate	95% Confidence Interval*
False Negatives	0.3670%	0.1001 to 0.9369%
False Negatives + inconclusive	1.376%	0.7722 to 2.260%
Poor mark production	2.319%	2.174 to 2.827%
False Positives	0.939% (1.010%)**	0.360% to 2.261%

* Confidence intervals for false negatives, false negatives plus inconclusive results for known same source samples, and poor mark production are calculated using the widely accepted Clopper-Pearson method for calculation of confidence intervals [7].
** As discussed in the text, we report a maximum likelihood estimator derived from a beta-binomial model with the raw proportion of errors in parentheses.

Calculations were made using the R statistical programming system (http://R-project.org). Clopper-Pearson confidence limits for the false-negative probability were computed using the binom.test routine; profile likelihood was used to compute the maximum likelihood estimate and confidence interval for the false-positive probability using the betabinom routine, which is in the VGAM library.

The distribution of unsuitable knowns was randomly distributed in the sample sets provided and did not affect either the false-negative or false-positive rate disproportionately. Of the 1090 known same-source sets analyzed, 66 (6.1%) included less than 3 suitable knowns. Of the 2178 known different-source sets analyzed, 148 (6.8%) included less than 3 suitable knowns. These are not significant differences. No correlation was observed between the occurrence of errors and the rate of poorly marked knowns by a particular firearm or group of firearms. Only two of the 26 errors (false-negative and -positive combined) corresponded with reports of less than 3 suitable knowns (in both cases the examiners reported 2 useable knowns).

Conclusions:

One of the goals of this project was determination of a false-negative rate for cartridge case comparisons. While we have a result based on the frequency of reporting of a finding of elimination for cartridge cases fired from the same handgun, this result remains in some doubt since this rate, even when combined with inconclusive findings from same-source comparisons, is smaller than the measured rate of production of fired cartridge cases unsuitable for comparison. Even if this poor pattern reproduction is not the cause of all of the reported false negatives, it does suggest that the rate we have measured is at best an upper limit on a true false-negative rate purely due to examiner error. One could design a study where the test materials are verified to be highly useable with very clear markings. However, the value of such a measurement is questionable. Since firearms in general are subject to this variability in mark production, if this phenomenon is coupled to the rate of reported false negatives, it may be significantly more important to study how variable this rate is within a particular model and between different models of firearms, instead of eliminating it from future study designs in order to measure a phenomenon (pure examiner error in finding a matching pattern) which is significantly less prevalent. Future studies with different firearms could be carried out with three or more of each model of several popular or frequently encountered firearms. The total number of examiners involved and amount of ammunition used might be much smaller than that used in this study and still be significant. A related study could also be designed to determine how different ammunition might affect the results for a subset of the study firearms. This might include cartridge cases in steel and brass, different manufacturers, and different batches in order to determine the significance of these variables.

Another goal of the project was to measure a false-positive rate for examiner cartridge case comparisons. This rate was measured and for the pool of participants used in this study the fraction of false positives was approximately 1%. The study was specifically designed to allow us to measure not simply a single number from a large number of comparisons, but also to provide statistical insight into the distribution and variability in false-positive error rates. The result is that we can tell that the overall fraction is not necessarily representative of a rate for each examiner in the pool. Instead, examination of the data shows that the rate is a highly heterogeneous mixture of a few examiners with higher rates and most examiners with much lower error rates. This finding does <u>not</u> mean that 1% of the time each examiner will make a false-positive error. Nor does it mean that 1% of the time laboratories or agencies would report false positives, since this study did not include standard or existing quality assurance procedures, such as peer review or blind reanalysis. What this result does suggest is that quality assurance is extremely important in firearms analysis and that an effective QA system must include the means to identify and correct issues with sufficient monitoring, proficiency testing, and checking in order to find false-positive errors that may be occurring at or below the rates observed in this study. Future research could be designed to determine the effectiveness of such quality assurance measures in finding and correcting these issues. If they involve peer review, these studies should examine potential effects like confirmation bias.

The final significant outcome of this study is the variability exhibited in the use of the inconclusive result among actively practicing firearms examiners. While it is entirely possible that examiners treated this study differently than they would casework, there are indications that the guidance of AFTE and SWGGUN is not uniformly followed or implemented by individuals or agencies. The rationale for the use of inconclusive is clear and justifiable as stated by AFTE. Some examiners or agencies choose not to use it at all and declare a comparison without sufficient detail for attribution to a common source to be an elimination. Others never used elimination during this study. Either of these alternatives may have been required by agency policy for these groups of participants. A large group of participants used both inconclusive and elimination as conclusions, indicating varying levels of certainty in the examinations. The wide variation in the use of these conclusions suggests a variation in interpretation that could weaken the basis of the range of conclusions as a standard. This issue is something that the discipline should address in order to clarify and strengthen the basis of all examiners' findings.

For the same-source samples, all four false-negative errors (false eliminations) were made by examiners who did not use inconclusive for any of their responses. This at least suggests that their agencies require that they declare that the comparison result in an identification or an elimination with no middle ground for poorly marked questioned samples, for which there is insufficient detail present to make such a judgment. Given the rate of unsuitable knowns reported for these same firearms, it is quite possible that these agencies are directing these examiners to declare an opinion about the source of casings based on samples with insufficient detail to make a definitive comparison. This can be likened to a fingerprint examiner being asked to identify an individual as either the same person or definitely not the same person who left a partial print at a scene of a crime when that print may be completely smudged or so fragmentary, with so little detail, that other reasonable examiners would say the print is unusable. This would not be a sound implementation of scientific analysis. The 11 examiners who reported inconclusive conclusions for the same-source samples have provided a result that is more descriptive of the evidence presented for poorly marked questioned samples. Although implementation of restrictions on conclusions is a matter of policy rather than of statistical analysis, it is entirely possible that no false-negative responses would have been received in this study had the range of conclusions used by all participants been reflective of the observations of the trained examiners involved instead of current policies. On the other hand, for the different-source samples, 735 were reported as inconclusive and 1421 were reported as eliminations. The fraction of samples reported as inconclusive cannot be attributed to a large fraction of poorly marked knowns or questioned samples in this group. Instead this fraction reflects the large number of examiners who through training or agency policy choose only to report inconclusive conclusions in the absence of an identification or class characteristics differences that would lead to an elimination. We suggest that it may be worth considering whether the use of inconclusive as a matter of policy on such a wide range of observations makes its use less effective in communicating exactly what is inconclusive about the analysis. Some examiners at a national AFTE meeting, when asked by one of the authors of this report about why they would use a mixture of inconclusive and elimination, suggested that policies that restrict conclusions to just identification or inconclusive are too restrictive and don't provide enough detail to report actual observations. Although this is once again a matter of policy rather than statistical significance, we suggest that sound scientific reporting would require conclusions describe observations and that policy makers would be better served by accurate analyses than by limited reporting. In the interests of standardization and clear communication, the means of expressing the findings may need to be uniform, but the present circumstances may be leading to misunderstandings about the meaning of inconclusive through its overuse.

The sample set created for this study, known as the DFSC Ruger SR9 Cartridge Case Set, is a valuable asset that will be used in future research. Not only is it a sizable collection of cartridge cases for which we know the source information regarding which firearm was used to generate each sample, but we have information from professional firearms examiners on the quality and detail of marks and comparisons with same-source and different-source samples. For individuals interested in using the sample set for research please email Henry.P.Maynard2.ctr@mail.mil or usarmy.gillem.dfsc.mbx.email@mail.mil. One of the first follow-up projects will involve collection of high quality 3D digital images of the sample to aid in the development of algorithms for computer comparisons and quantification of the amount of detail available for comparisons. This information could be used for evaluation of human response thresholds and the development of software tools to aid in comparisons.

References:

1) J. Gouwe, J. Hamby, and S. Norris, "Comparison of 10,000 Consecutively Fired Cartridge Cases from a Model 22 Glock .40 S&W Caliber Semiautomatic Pistol," *AFTE Journal* **40** (1), 57-63 (2008).

2) Bunch, S. G., Murphy D., "A Comprehensive Validity Study for the Forensic Examination of Cartridge Cases," *AFTE Journal* **35** (2), 201-203 (2003).

3) James Hamby and James Thorpe, "The Examination, Evaluation and Identification of 9mm Cartridge Cases Fired from 617 Different Glock Model 17 & 19 Semiautomatic Pistols", *AFTE Journal* **41** (4) 310-324 (2009).

4) LaPorte, D., "An Empirical Validation Study of Breechface Marks on .380 ACP Caliber Cartridge Cases Fired from Ten Consecutively Finished Hi-Point Model C9 Pistols," *AFTE Journal* **43** (4), (2011).

5) "An Empirical Study to Improve the Scientific Foundation of Forensic Firearm and Tool Mark Identification Utilizing 10 Consecutively Manufactured Slides", Award Number 2009-DN-BX-K230, Final Report, Submitted By: Miami-Dade Police Department Crime Laboratory, 2011. https://www.ncjrs.gov/pdffiles1/nij/grants/237960.pdf

6) Committee on Identifying the Needs of the Forensic Sciences Community, National Research Council, "Strengthening Forensic Science in the United States: A Path Forward," National Academies Press, p. 154 (2009).

7) Clopper, C. J., and Pearson, E. S., "The Use of Confidence or Fiducial Limits Illustrated in the Case of the Binomial," *Biometrika*, 26, 404-413 (1934).

Appendix A

Invitation and Informed Consent

THE Ames Laboratory
Creating Materials & Energy Solutions
U.S. DEPARTMENT OF ENERGY

June 14, 2013

Dear Firearm Examiner,

You are being invited to participate in a reliability study evaluating false positive and false negative errors in comparing cartridge casings. Only firearm examiners who currently conduct examinations and are members of AFTE or are employed in the firearms section of an ASCLD member crime laboratory are being asked to participate. This study is sponsored by the Defense Forensic Office and will be carried out by the Midwest Forensics Resource Center (MFRC) at the Ames Laboratory USDOE.

Participation is completely voluntary. There is no compensation for participating in this study. Participating examiners will be sent a set of cartridge casings and asked to compare known and questioned casings. Through a survey instrument, you will be asked to conclude whether the compared casings are identifications, inconclusive, eliminations, or unsuitable. Reported results and findings will be completely anonymized. Additional information about experience, certification, lab accreditation, method, and instrumentation will also be collected through the survey instrument. The study findings will result in a peer-reviewed publication that will be relevant to the legal admissibility of such analyzes.

If you are interested in participating in this study, please complete the attached consent form and return to the MFRC at mfrc@ameslab.gov or by mail (Attn: Firearms Study, 127 Spedding Hall, Ames, IA 50011-3020) by August 31, 2013. If you need further information, please contact Dr. David Baldwin – MFRC Director ([redacted]).

Sincerely,

[redacted]

David P. Baldwin

IOWA STATE UNIVERSITY
OF SCIENCE AND TECHNOLOGY

Ames Laboratory is operated by Iowa State University for the U.S. Department of Energy, Ames, Iowa 50011-3020

CONSENT FORM FOR: Investigation of False Positive and False Negative Errors for Cartridge Casing Comparisons

This form describes a research project. It has information to help you decide whether or not you wish to participate. Research studies include only people who choose to take part—your participation is completely voluntary. Please discuss any questions you have about the study or about this form with the project staff before deciding to participate.

Who is conducting this study?

This study is being conducted by Dr. David P. Baldwin from the Midwest Forensics Resource Center (MFRC) at the Ames Laboratory USDOE. This study is funded by the Defense Forensic Office.

Why am I invited to participate in this study?

You are being asked to take part in this study because you are either an Association of Firearm and Tool Mark Examiners (AFTE) member or perform firearm examinations at an ASCLD member crime laboratory. You should not participate if you are not currently performing firearm examinations as part of your normal employment duties.

What is the purpose of this study?

The purpose of this study is to evaluate the reliability of firearms examiners in the analysis and comparison of cartridge casings in order to determine error rates and the degree of correlation between these rates and various related factors.

What will I be asked to do?

If you agree to participate, you will be sent a set of cartridge casings and asked to compare known and questioned casings. This set will contain fifteen (15) sets of cartridge casings for comparison. Through a survey instrument, you will be asked to conclude whether the compared casings are identifications, inconclusive, eliminations, or unsuitable. For microscopic examination, reported results and findings will be completely anonymized. Additional information about experience, certification, lab accreditation, method, and instrumentation will also be collected through the survey instrument.

Your participation will last for the length of time it takes to examine the sets of casings provided and to fill out an answer sheet and survey.

What are the possible risks and benefits of my participation?

Risks—there are no known risks related to your participation in this research.

Benefits—you may not receive any direct benefit from taking part in this study. We hope that this research will benefit society by providing a better statistical evaluation of this common and important forensic discipline that will strengthen the legal system in its understanding of the value of firearms comparisons.

How will the information I provide be used?

The information you provide will be used for the following purposes: Perform a statistical analysis in order to determine error rates and degree of correlation with these rates with various related factors.

What measures will be taken to ensure the confidentiality of the data or to protect my privacy?

Records identifying participants will be kept confidential to the extent allowed by applicable laws and regulations. Records will not be made publicly available. However, federal government regulatory agencies [DOE, DoD], auditing departments of Iowa State University, and the ISU Institutional Review Board (a committee that reviews and approves research studies with human subjects) may inspect and/or copy your records for quality assurance and analysis. These records may contain private information. However, at no point will your survey responses or the results of your examinations be connected to your identity. Only the fact that you participated may be revealed by these reviews and audits. Your participation is only known via this signed consent form. Your identity will remain confidential if these results are published.

To ensure confidentiality to the extent allowed by law, the following measures will be taken: 1) Participant contact information will be kept on a password-protected computer and only accessed by administrative staff; 2) Cartridge casing sample set identifiers will not be linked to any participant; 3) Survey instruments will not contain any identifiable information and will be stored in a locked filing cabinet and access limited to the researchers of this study. This information will be kept for three years after completion of the project. If the results are published, your identity will remain confidential.

Will I incur any costs from participating or will I be compensated?

You will not have any costs from participating in this study. You will not be compensated for participating in this study.

What are my rights as a human research participant?

Participating in this study is completely voluntary. You may choose not to take part in the study or to stop participating at any time, for any reason, without penalty or negative consequences. You can skip any questions that you do not wish to answer.

Whom can I call if I have questions or problems?

You are encouraged to ask questions at any time during this study.

- For further information about the study contact Dr. David Baldwin – MFRC Director (b)(6) per EOUSA [redacted]).
- If you have any questions about the rights of research subjects or research-related injury, please contact the IRB Administrator, (515) 294-4566, IRB@iastate.edu, or Director, (515) 294-3115, Office for Responsible Research, 1138 Pearson Hall, Iowa State University, Ames, Iowa 50011.

Consent and Authorization Provisions

Your signature indicates that you voluntarily agree to participate in this study, that the study has been explained to you, that you have been given the time to read the document and that your questions have been satisfactorily answered. You will receive a copy of the written informed consent prior to your participation in the study.

Additionally, your signature indicates that upon completion of your participation in this study, you agree not to discuss with other examiners that may be participating details about this study or your findings, so that their contribution and findings may be unbiased and independent.

Participant's Name (printed) ______________________________

Participant's Phone No.______________ Participant's email____________________

______________________________ ______________________
(Participant's Signature) (Date)

______________________________ ______________________
(Signature of Lab Director or Section Supervisor *required only when applicable or necessary*) (Date)

Shipping Address to receive study materials (UPS – No P.O. Box addresses, please)

____

____

____

Please return signed form to mfrc@ameslab.gov or mail to:

MFRC, Attn: Firearms
127 Spedding Hall
Ames, IA, 50011-3020

Appendix B

Table of Firearm Serial Numbers and Ammunition Batch Information

Gun	Letter Designation	Serial Number	Ammo Lot#		
1	A1	331-96383	3503 ODFL	2195980	
2	A2	331-96385	3503 ODFL	2195980	
3	A3	331-96387	3503 ODFL	2195980	
4	A4	331-96388	3503 ODFL	2195980	
5	A5	331-96584	3503 ODFL	2195980	
6	B1	331-96585	3503 ODFL	2195980	
7	B2	331-96586	3503 ODFL	2195980	
8	B3	331-96590	3503 ODFL	2195980	
9	B4	331-96592	3503 ODFL	2195980	(1st 400)
			350 ODFL	32206416	(2nd 400)
10	B5	331-96593	350 ODFL	32206416	
11	C1	331-96594	350 ODFL	32206416	
12	C2	331-96604	350 ODFL	32206416	
13	C3	331-96620	350 ODFL	32206416	
14	C4	331-96649	350 ODFL	32206416	
15	C5	331-96651	350 ODFL	32206416	
16	D1	331-96661	350 ODFL	32206416	
17	D2	331-96663	350 ODFL	32206416	
18	D3	331-96664	350 ODFL	32206416	
19	D4	331-96665	350 ODFL	32206416	
20	D5	331-96667	350 ODFL	32206416	
21	E1	331-96669	350 ODFL	32206416	
22	E2	331-96681	350 ODFL	32206416	
23	E3	331-96689	350 ODFL	32206416	
24	E4	331-96718	350 ODFL	32206416	
25	E5	331-96719	350 ODFL	32206416	

Appendix C

Survey and Answer Form

Comparison Group No. ____________________

SURVEY QUESTIONS:

Years Experience:______ Years Training:______

AFTE Certified: Yes☐ No☐ ABC Certified: Yes☐ No☐ Other: ______________

Attended the FBI Specialized Techniques School: Yes☐ No☐ CMS Trained: Yes☐ No☐

Do you work in a firearms ASCLD-member laboratory: Yes☐ No☐

Do you currently conduct firearms casework: Yes☐ No☐

Do you examine other types of evidence: Yes☐ No☐ If Yes, what other types __________

____

Brand & Model of Microscope used:______________________________

Type of Lighting used:______________________________

Did you use Pattern Matching, CMS or Both for this test? ______________

Work performed in accredited laboratory: Yes☐ No☐ Are you an AFTE member: Yes☐ No☐

Set No. Number of known's with sufficient reproduced detail for comparison: 0☐ 1☐ 2☐ 3☐

1 Identification ☐ Elimination ☐ Inconclusive (Please provide basis) Unsuitable ☐

a) ☐ Some agreement of individual characteristics and all discernible class characteristics, but insufficient for an identification.
b) ☐ Agreement of all discernible class characteristics without agreement or disagreement of individual characteristics due to an absence, insufficiency, or lack of reproducibility.
c) ☐ Agreement of all discernible class characteristics and disagreement of individual characteristics, but insufficient for an elimination.

Set No. Number of known's with sufficient reproduced detail for comparison: 0☐ 1☐ 2☐ 3☐

2 Identification ☐ Elimination ☐ Inconclusive (Please provide basis) Unsuitable ☐

a) ☐ Some agreement of individual characteristics and all discernible class characteristics, but insufficient for an identification.
b) ☐ Agreement of all discernible class characteristics without agreement or disagreement of individual characteristics due to an absence, insufficiency, or lack of reproducibility.
c) ☐ Agreement of all discernible class characteristics and disagreement of individual characteristics, but insufficient for an elimination.

Set No. Number of known's with sufficient reproduced detail for comparison: 0☐ 1☐ 2☐ 3☐

3 Identification ☐ Elimination ☐ Inconclusive (Please provide basis) Unsuitable ☐

a) ☐ Some agreement of individual characteristics and all discernible class characteristics, but insufficient for an identification.
b) ☐ Agreement of all discernible class characteristics without agreement or disagreement of individual characteristics due to an absence, insufficiency, or lack of reproducibility.
c) ☐ Agreement of all discernible class characteristics and disagreement of individual characteristics, but insufficient for an elimination.

(over)

Set No. Number of known's with sufficient reproduced detail for comparison: 0☐ 1☐ 2☐ 3☐

4 Identification ☐ Elimination ☐ Inconclusive (Please provide basis) Unsuitable ☐

a) ☐ Some agreement of individual characteristics and all discernible class characteristics, but insufficient for an identification.
b) ☐ Agreement of all discernible class characteristics without agreement or disagreement of individual characteristics due to an absence, insufficiency, or lack of reproducibility.
c) ☐ Agreement of all discernible class characteristics and disagreement of individual characteristics, but insufficient for an elimination.

Set No. Number of known's with sufficient reproduced detail for comparison: 0☐ 1☐ 2☐ 3☐

5 Identification ☐ Elimination ☐ Inconclusive (Please provide basis) Unsuitable ☐

a) ☐ Some agreement of individual characteristics and all discernible class characteristics, but insufficient for an identification.
b) ☐ Agreement of all discernible class characteristics without agreement or disagreement of individual characteristics due to an absence, insufficiency, or lack of reproducibility.
c) ☐ Agreement of all discernible class characteristics and disagreement of individual characteristics, but insufficient for an elimination.

Set No. Number of known's with sufficient reproduced detail for comparison: 0☐ 1☐ 2☐ 3☐

6 Identification ☐ Elimination ☐ Inconclusive (Please provide basis) Unsuitable ☐

a) ☐ Some agreement of individual characteristics and all discernible class characteristics, but insufficient for an identification.
b) ☐ Agreement of all discernible class characteristics without agreement or disagreement of individual characteristics due to an absence, insufficiency, or lack of reproducibility.
c) ☐ Agreement of all discernible class characteristics and disagreement of individual characteristics, but insufficient for an elimination.

Set No. Number of known's with sufficient reproduced detail for comparison: 0☐ 1☐ 2☐ 3☐

7 Identification ☐ Elimination ☐ Inconclusive (Please provide basis) Unsuitable ☐

a) ☐ Some agreement of individual characteristics and all discernible class characteristics, but insufficient for an identification.
b) ☐ Agreement of all discernible class characteristics without agreement or disagreement of individual characteristics due to an absence, insufficiency, or lack of reproducibility.
c) ☐ Agreement of all discernible class characteristics and disagreement of individual characteristics, but insufficient for an elimination.

Set No. Number of known's with sufficient reproduced detail for comparison: 0☐ 1☐ 2☐ 3☐

8 Identification ☐ Elimination ☐ Inconclusive (Please provide basis) ☐

a) ☐ Some agreement of individual characteristics and all discernible class characteristics, but insufficient for an identification.
b) ☐ Agreement of all discernible class characteristics without agreement or disagreement of individual characteristics due to an absence, insufficiency, or lack of reproducibility.
c) ☐ Agreement of all discernible class characteristics and disagreement of individual characteristics, but insufficient for an elimination.

Set No. Number of known's with sufficient reproduced detail for comparison: 0☐ 1☐ 2☐ 3☐

9 Identification ☐ Elimination ☐ Inconclusive (Please provide basis) Unsuitable ☐

a) ☐ Some agreement of individual characteristics and all discernible class characteristics, but insufficient for an identification.

(next page)

b) ☐ Agreement of all discernible class characteristics without agreement or disagreement of individual characteristics due to an absence, insufficiency, or lack of reproducibility.
c) ☐ Agreement of all discernible class characteristics and disagreement of individual characteristics, but insufficient for an elimination.

Set No. Number of known's with sufficient reproduced detail for comparison: 0 ☐ 1 ☐ 2 ☐ 3 ☐

10 Identification ☐ Elimination ☐ Inconclusive (Please provide basis) Unsuitable ☐

a) ☐ Some agreement of individual characteristics and all discernible class characteristics, but insufficient for an identification.
b) ☐ Agreement of all discernible class characteristics without agreement or disagreement of individual characteristics due to an absence, insufficiency, or lack of reproducibility.
c) ☐ Agreement of all discernible class characteristics and disagreement of individual characteristics, but insufficient for an elimination.

Set No. Number of known's with sufficient reproduced detail for comparison: 0 ☐ 1 ☐ 2 ☐ 3 ☐

11 Identification ☐ Elimination ☐ Inconclusive (Please provide basis) Unsuitable ☐

a) ☐ Some agreement of individual characteristics and all discernible class characteristics, but insufficient for an identification.
b) ☐ Agreement of all discernible class characteristics without agreement or disagreement of individual characteristics due to an absence, insufficiency, or lack of reproducibility.
c) ☐ Agreement of all discernible class characteristics and disagreement of individual characteristics, but insufficient for an elimination.

Set No. Number of known's with sufficient reproduced detail for comparison: 0 ☐ 1 ☐ 2 ☐ 3 ☐

12 Identification ☐ Elimination ☐ Inconclusive (Please provide basis) Unsuitable ☐

a) ☐ Some agreement of individual characteristics and all discernible class characteristics, but insufficient for an identification.
b) ☐ Agreement of all discernible class characteristics without agreement or disagreement of individual characteristics due to an absence, insufficiency, or lack of reproducibility.
c) ☐ Agreement of all discernible class characteristics and disagreement of individual characteristics, but insufficient for an elimination.

Set No. Number of known's with sufficient reproduced detail for comparison: 0 ☐ 1 ☐ 2 ☐ 3 ☐

13 Identification ☐ Elimination ☐ Inconclusive (Please provide basis) Unsuitable ☐

a) ☐ Some agreement of individual characteristics and all discernible class characteristics, but insufficient for an identification.
b) ☐ Agreement of all discernible class characteristics without agreement or disagreement of individual characteristics due to an absence, insufficiency, or lack of reproducibility.
c) ☐ Agreement of all discernible class characteristics and disagreement of individual characteristics, but insufficient for an elimination.

Set No. Number of known's with sufficient reproduced detail for comparison: 0 ☐ 1 ☐ 2 ☐ 3 ☐

14 Identification ☐ Elimination ☐ Inconclusive (Please provide basis) Unsuitable ☐

a) ☐ Some agreement of individual characteristics and all discernible class characteristics, but insufficient for an identification.
b) ☐ Agreement of all discernible class characteristics without agreement or disagreement of individual characteristics due to an absence, insufficiency, or lack of reproducibility.
c) ☐ Agreement of all discernible class characteristics and disagreement

---

(over)

of individual characteristics, but insufficient for an elimination.

Set No.	Number of known's with sufficient reproduced detail for comparison: 0 ☐ 1 ☐ 2 ☐ 3 ☐			
15	Identification	Elimination	Inconclusive (Please provide basis)	Unsuitable
	☐	☐	a) ☐ Some agreement of individual characteristics and all discernible class characteristics, but insufficient for an identification. b) ☐ Agreement of all discernible class characteristics without agreement or disagreement of individual characteristics due to an absence, insufficiency, or lack of reproducibility. c) ☐ Agreement of all discernible class characteristics and disagreement of individual characteristics, but insufficient for an elimination.	☐

(next page)

Appendix D

Participant Instructions

INSTRUCTION SHEET

ITEMS CONTAINED IN THIS PACKAGE:

-This instruction sheet.
-A sealed and unmarked envelope containing the cartridge cases for comparison, a survey, and a recording sheet to record your findings on.
-An unsealed and labeled UPS Express shipping box and an unsealed and unmarked envelop to return the cartridge cases, survey, and recording sheet.

HOW TO CONDUCT THE COMPARISONS:

The cartridge cases are divided into sets of three known's (k's –produced from the same firearm) and a questioned (q). Each *set* comes in its own numbered envelope. This number corresponds to the set number on the recording sheet. Each *set* envelop contains two envelopes which contain the k's and q respectively.

Work with one set at a time. There are a total of 15 sets. Return cartridge cases to their respective envelopes/bags after comparison is made in order to minimize mixing up the casings. Record findings for each set on provided sheet.

You are being asked to compare the k's to the q and render a finding of either "Identification, Elimination, Inconclusive, or Unsuitable," as defined in the AFTE Glossary (Appendix 1) "Range of Conclusions Possible when Comparing Toolmarks."

Indicate on the recording sheet the number of k's (none, one, two, or three) that have adequately reproduced marks for comparison and are adequate to confirm an identification between the k's from the known same source.

Finally, if your finding for a particular set is Inconclusive, please select a basis for this finding. The three choices come from the AFTE Glossary (Appendix 1) "Range of Conclusions Possible when Comparing Toolmarks."

NOTICE:

The purpose of this study is to measure error rates for individual examiners. Please do not peer-review or confirm results of your examination.

AFTER COMPARISONS ARE MADE:

Complete survey. Place survey, recording sheet, and cartridge case set in the provided unmarked envelope and seal. Please do not include any indentifying marks that would indicate your identity on either the survey or envelope.

Place sealed envelope (containing materials) into UPS Express shipping box and seal the box. Return materials to MFRC.

Please do *not* discuss your findings or this study with other examiners who may be participating, so that their contribution and findings may be unbiased and independent.

++

If you have any questions, please contact either Ms. Melinda Schlosser (515-296-6372; mschlosser@ameslab.gov) or Dr. Rudi Luyendijk (515-294-2931; rluyendi@ameslab.gov) for assistance. Thank you for your participation!

JOURNAL OF FORENSIC SCIENCES

J Forensic Sci, 2016
doi: 10.1111/1556 4029.13093
Available online at: onlinelibrary.wiley.com

PAPER

CRIMINALISTICS

Tasha P. Smith,[1] *B.S.; G. Andrew Smith,*[1] *M.S.; and Jeffrey B. Snipes,*[2] *J.D., Ph.D.*

A Validation Study of Bullet and Cartridge Case Comparisons Using Samples Representative of Actual Casework

ABSTRACT: The foundation of firearm and tool mark identification is that no two tools should produce the same microscopic marks on two separate objects that they would be inaccurately or wrongly identified. Studies addressing the validity of identification infrequently employ tests that mirror realistic casework scenarios. This study attempted to do so using a double blind process, reducing test taking bias. Test kits including bullets and cartridge cases but not the associated firearms were completed by 31 analysts from 22 agencies. Analysis of the results demonstrated an overall error rate of 0.303%, sensitivity of 85.2%, and specificity of 86.8%. Variability in performance across examiners is addressed, and the effect of examiners' years of experience on identification accuracy is explored. Finally, the article discusses the importance of studies using realistic case work scenarios when validating the field's performance and in providing courts with usable indicators of the accuracy of firearm and tool mark identification.

KEYWORDS: forensic science, firearm and tool mark identification, error rate studies, validation study, casework simulation, sensitivity and specificity

The foundation of the science of firearm and tool mark identi fication is that no two tools should produce the same micro scopic marks on two separate objects that they would be inaccurately or wrongly identified. Firearm identification relies upon the human cognitive ability of pattern recognition that allows one to determine the individuality of a tool, through the physical comparison of microscopic marks. Years of research have proven that in the evaluation of consecutively manufactured tools tools which show the greatest potential for leaving the same marks these tools display sufficient individual differences that when subclass influence is excused, the origin of marks left by consecutively manufactured tools can be determined. The firearm and tool mark examiner often faces several common core questions, such as: Is it possible to identify or exclude a tool as having created a mark from all other possible tools? Can such exclusions and identifications be made with any degree of cer tainty? What is the range of certainty of this exclusion or identi fication? How do these findings translate to the everyday community or courts in way that is easy to understand by the layperson?

Much of the research to date has supported the theory of indi vidualization and has been performed so through the microscopic comparison and observation of barrels, slides, knife blades, screwdrivers, and so forth (1 5). Some research has further been complemented by the use of statistical and mathematical models (6 12). Yet, often the validity of these measures is criticized (13 15).

While it is true that errors occur in all human endeavors, whether in computer programming or in an emergency room, the crucial benchmark for bases of comparison over time or across agencies/organizations is the frequency and likelihood of occur rence of these errors. In firearm and tool mark identification, the frequency with which errors occur is difficult to deduce because the outcome of the work is dependent on the presence of con trols and quality checks. With mounting methodological criti cisms and case decisions, the courts are not interested in a "theoretical error rate," which assumes that everything has been carried out properly and the correct answers have been reached. What they are interested in and what is of more value is what actually happens during routine casework. Additionally, courts want this data be reported with a level of understanding, cer tainty, and specificity of that commonly seen in DNA analyses (14). However, the level of understanding in firearm and tool mark identification that corresponds to that level of DNA analy sis exists only on a subclass level, not on an individual level. The "human factor" in identification accounts for tremendous variability in analysis. Some of the most important questions that have arisen with validation studies include as follows:

- Can a validation study which is representative of actual case work in the field of firearm and tool mark examination be designed and implemented?
- Can this study be presented in a blind or double blind format?
- Can such a test be designed that addresses the possibility of test taking bias?
- Can the results be tabulated with a level of accuracy that is reasonably consistent across all examiners?

[1]San Francisco Police, Department Crime Lab, 1245 3rd Street, San Francisco, CA, 94158.
[2]Program in Criminal Justice Studies, San Francisco State University, San Francisco, CA.

Received 8 April 2015; and in revised form 18 Sept. 2015; accepted 4 Oct. 2015.

- To what extent is training and experience a factor in the examiner's decision making process and outcome?
- Are results and error rate values consistent across studies and are they representative of actual casework values?
- Can these results be articulated in a way that is understand able and of value to the community in a precisely specified and scientifically justified way that leads to a well character ized confidence limit?

The training of a firearms examiner is based on the understand ing of the individualizing marks produced, where they come from and how they are made. This training involves a constant build ing and refining of what is called an examiner's criteria for iden tification. The criteria for identification are a subjective point refined through the experience and training of an examiner of what is sufficient and significant agreement in the individual microscopic marks of interest. Such a level of understanding can not always be conveyed quantitatively; however, through meth ods such as QCMS the level of agreement that can be translated in a fashion understandable to the general public is approachable. Quantitative consecutive matching stria (QCMS) is a method of identification that provides a quantitative value to the evaluation of striated marks. Although QCMS is becoming more widely used in the field of firearm and tool mark identification, it is lim ited in that it only applies to striated marks. It is also limited in determining which lines in a pattern can be counted versus those that should not. When solely using pattern matching, it is the combination of the overall similarity of the pattern and the micro scopic detail of the pattern of both striated and impressed marks that must meet an examiner's criteria for identification for an identification to be made. An examiner's knowledge base can only be developed and refined through the constant and consis tent evaluation of known matches (KMs) to known nonmatches (KNMs) that allow for the assessment of individuality.

The purpose of this study is to present the design and results of a study that has been developed to provide the discipline with a useable accurate error rate that is a clear and concise represen tation of the actual human work associated with firearms tool mark identification. It also addresses variability in sensitivity and specificity measures across multiple examiners. Finally, it attempts to determine whether there is any relationship between an examiner's years of experience and performance in identifica tion.

Materials and Methods

Test Design

Each test was designed to have a similar feel to what an exam iner typically encounters when working a case. It is routine within a criminalistics laboratory that a firearm examiner will receive evidence with little knowledge of the history of the evidence and such evidence is often presented without a firearm. Such situa tions limit what examiners have to make comparisons with, while also testing their knowledge of manufacturing processes, what is possible, and what is probable, in the operation of firearms. This study aimed to approximate everyday casework by providing examiners with a realistic, albeit simulated, case with no firearm. Such a design should provide a more realistic assessment of error rates in case work. This study is similar to a number of other studies; however, there are marked differences in the design to make it more realistic to what is seen on the bench on a daily basis. Like the studies by Smith (16) and others, the firearms used for test firing were obtained from crime related cases and there fore were circulated in the general population and subjected to use, corrosion and abuse similar to that observed in a typical case. These tests were then circulated to active firearms examiners with varying years of experience and levels of training, working in lab oratories which vary in their policies and procedures for making exclusions when the firearm is absent.

A primary criticism of many of the reported validation studies within the community is that many tests lack anonymity and some examiners are more conservative than others due to the fear of answering incorrectly. This may create a test taking bias. The current test was as blind as possible except to the extent participants were aware that they were participating in a valida tion study. To provide as much separation as possible between researcher and participants, requests for participants were sent out by a third party via email or message board to maximize sampling randomness and eliminate any questions of bias between test administrator and the participants. All test takers and supervisors were unaware of the correct answers, and the test administrator was not privy to which individual in a particu lar laboratory was taking the test. Each test packet was different from the next, eliminating the likelihood of discussions between participants within the same laboratory resulting in any useful information being obtained. Although a number of the tests were sent out multiple times OR sent out on multiple occasions, they were never duplicated within the same laboratory. This not only provided us with a measure of reproducibility but also served as a quality check of the tests themselves. Each test was of similar difficulty. The number of identifications to exclusions varied from test to test, containing anywhere from 12 to 14 true identi fications and 20 30 true eliminations as designed.

This study utilized both bullets and cartridge cases from eight different firearms that had been circulated in the general popula tion and now reside in the San Francisco Police Department Crime Laboratory's Firearm Reference Collection. These fire arms consisted of at least two with the same class characteristics; therefore, an evaluation of individual microscopic marks was necessary. A total of 406 true identifications and 760 true elimi nations were possible within the 31 returned kits as they were designed. There were 1060 actual eliminations possible based on the "if then" result of the actual conclusions within the test. The number of possible eliminations to identifications sought to chal lenge the examiners' criteria for identification using either pat tern recognition or quantitative consecutive matching striae criteria while also challenging any testing preconceptions devel oped through the participation in other similar studies. In this study, there were no "knowns" with which to compare "un knowns." This feature is not usually found in traditional studies but is more reflective of the actual level of comparison work that an examiner may encounter. All test sets in this study consisted of at least one cartridge case and/or bullet (or bullet jacket) that did not identify to any other specimen within the test kit.

Materials

Six different types of ammunition consisting of 1104 car tridges were fired through eight different 40 caliber pistols. The various firearms were used because of their unique ability to mark ammunition in ways consistent with what is seen in every day casework. Two different firearms of a similar make and model for each of the four firearm types were used. The make, model, general rifling characteristics, and serial numbers of the firearms used in this study are documented in Table 1.

TABLE 1 *Types of firearms used from SFPD reference collection.*

Make	Model	Caliber	GRC	Serial Number	Ammunition Type Fired per Firearm
Taurus	PT 101 AFS	0.40	6R	SLD18629D	92 UMC (CC and Bu); 92 WIN BEB (CC and Bu); 92 Hi Shok/Hydra shok (Bu); 92 American Eagle (CC)
	PT 101 AF	0.40	6R	SKJ01550/AFD	
Sig Sauer	P229	0.40	6R	AC19988	92 UMC (CC and Bu); 92 Speer GD (CC and Bu); 92 Hi Shok/Hydra Shok (Bu); 92 American Eagle (CC)
	P229	0.40	6R	AC16713	
Smith and Wesson	4013	0.40	6L	THZ9553	92 UMC (Bu); 92 WIN BEB (Bu); 92 Hi Shok/Hydra shok (Bu)
	SW40C	0.40	6L	PAL5819	
Glock	22	0.40	6R	ARC775US	92 UMC (CC); 92 WIN BEB (CC); 92 American Eagle (CC)
	27	0.40	6R	CZR349US	

GRC, General Rifling Characteristics; CC, Cartridge cases; Bu, Bullets.

Six different types of ammunition were used in the execution of this study. A list of ammunition specifications is found in Table 2. Each of the fired bullets and cartridge cases was assigned a unique identifying number as a key. To decrease the chances of a recognizable pattern being observed by test takers, the identifying numbers were obtained using a random number generator program (17). The identifying number was inscribed on the ogive or base of the bullet and jackets; and on the side of the cartridge case using a Dremel model 290 01 engraver. The cartridges were fired into a horizontal water tank equipped with a "lab made" bullet retrieval trap, which was constructed using PVC pipe cut to the dimensions of the tank with durable mesh screen along the bottom. The design and use of this trap allowed for rapid collection of the multiple specimens fired in this study. Representative samples of some of the specimens from the test are provided in Figs 1 6.

Packet Preparation

A total of 50 study packets were prepared, each containing 12 randomly selected bullets/bullet jackets and 12 randomly selected cartridge cases, a supplementary comparison worksheet, an answer sheet, and directions for performing the study (Appendix S1). Each test packet was given its own unique iden tifier to maintain anonymity of the test participants. Participating laboratories were sent 1 3 packets at their request that were dis tributed by the supervisor, in most cases, to bench level analysts. A total of 47 kits were distributed, with 34 returned, three of which were omitted because they violated the conditions of the study in one way or another.

For the purposes of collection, each firearm was fired individu ally, with all the specimens collected and placed into individually labeled containers. The container was labeled with the firearm make, model, and serial number information. The specimens were later engraved with a unique identifier supplied through the ran dom number generator program. It should be noted that in some cases, only bullets/bullet jackets were collected, such as for the Sig Sauer and Smith and Wesson firearms. And in some cases, only the cartridge cases were collected, such as with the Glock firearms. The total evidence specimen count was 2208, of which 1200 were placed into 50 kits (containing 12 bullets/bullet jackets and 12 cartridge cases). The randomness of this study was maxi mized by thoroughly mixing all of the bullets/jackets after being scribed with their identifiers. Then, 12 were randomly selected and grouped from the container of bullets and cartridge cases by indi viduals from the laboratory. The scribed numbers were then recorded onto individual 2 ½" × 4 ¼" size envelopes and placed into the corresponding envelopes sealed with tape and then placed into individual test packets, labeled with a test number 1 thru 50. Over the next several days, examiners from the San Francisco Police Department Crime Lab Firearm and Tool Mark Unit evalu ated the kits for their potential for identification, and to ensure that where identifications should be made, they could be made. The examiners had a range of training histories and levels of experi ence, as did members of the testing group. Following the kit evalu ations, the test packets were sealed and shipped to the 47 participants representing approximately 30 different laboratories across the United States and abroad. Participants were given

TABLE 2 *Ammunition specifications.*

Ammunition Name/Brand	Cartridge	Grain	Primer	Case	Bullet Type/ Composition
Remington UMC	40 S&W	165/185	Nickel	Brass	FMJ/Copper
Federal Classic Hi Shok	40 S&W	155/180	Brass	Brass	JHP/Copper
Federal Classic Hydra Shok	40 S&W	155	Nickel	Nickel	JHP/Copper
Winchester WinClean BEB	40 S&W	165	Nickel	Brass	FMJ/Brass
Speer Gold Dot	40 S&W	180	Nickel	Nickel	JHP/Copper
American Eagle	40 S&W	180	Brass	Brass	FMJ/Copper

FMJ, Full Metal Jacket; JHP, Jacketed Hollow Point.

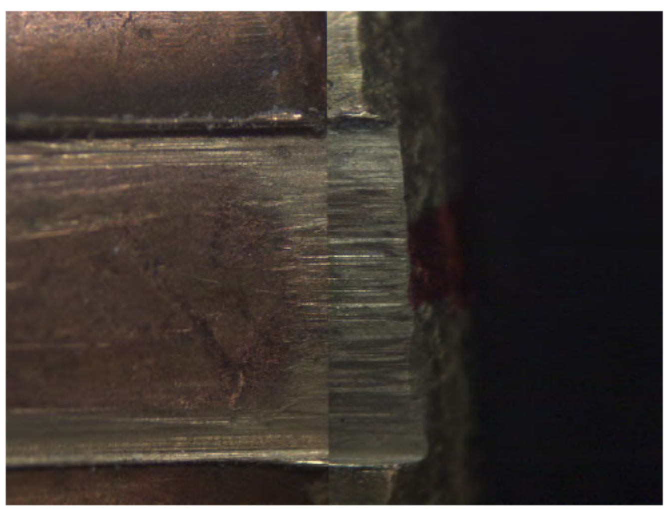


FIG. 1 *Kit # 22 Ex 1098 to Ex 1267 28X, LIMP 1.*[1]

[1]Land Impression

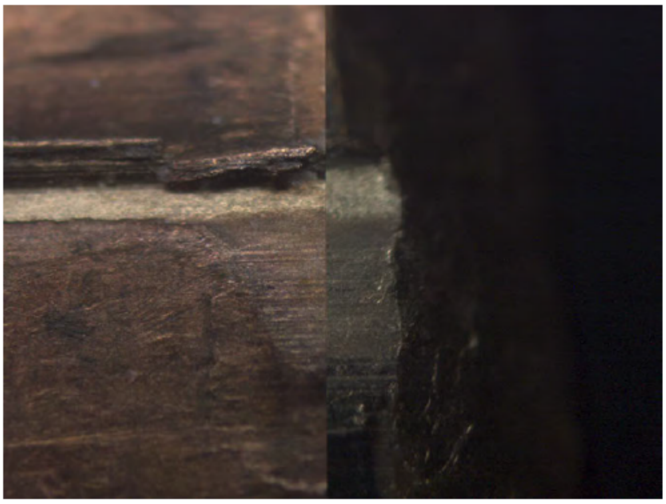


FIG. 2 *Kit # 22 Ex 1288 to Ex 1124 55X, LIMP 2.*

FIG. 3 *Kit # 22 Ex 1288 to Ex 1844 55X, LIMP 1.*

varying amounts of time to complete the test, based on phases of this research project, and it was requested that both answer sheets and kits be returned upon completion. Time duration was estimated to be between 2 and 12 months. Twenty two different laboratories/laboratory systems across the country (and one abroad) were represented in the results received.

Results

We report two types of analyses in this section. First, we examine the overall error rates, sensitivity and specificity levels, in an aggregate fashion with no attention given to differences in examiners. Second, we provide additional analysis that looks at sensitivity and specificity levels as they are distributed across the 31 examiners, as well as the effect of years of experience on identification performance.

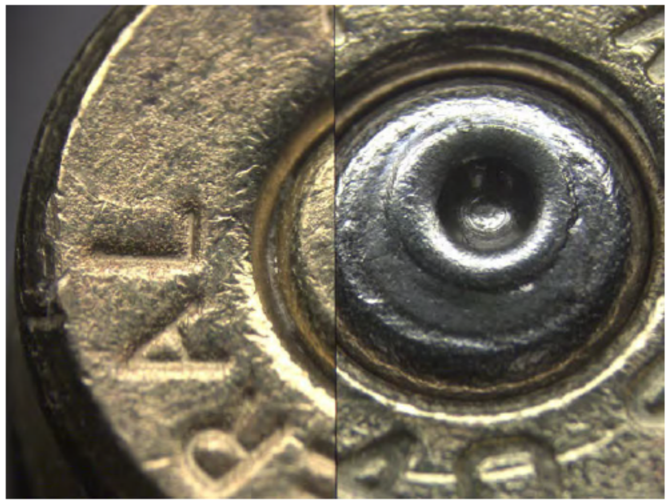


FIG. 4 *Kit # 27 Ex 1191 to Ex 1834 14X, BFM 2.*[2]

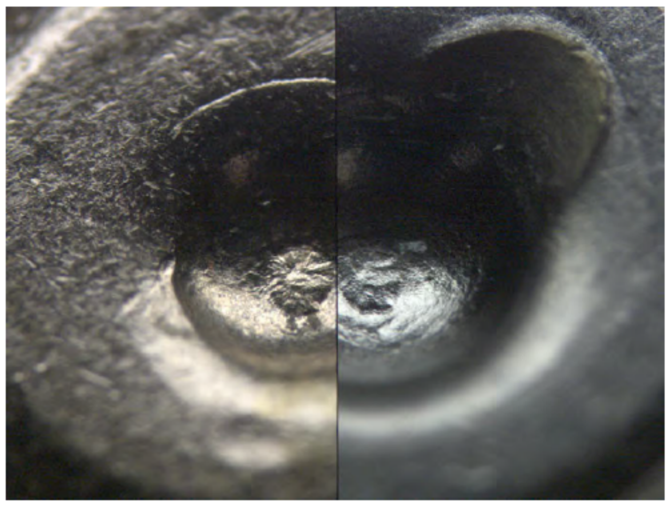


FIG. 5 *Kit #27 Ex 1238 to Ex 1760 35X, FPIM.*[3]

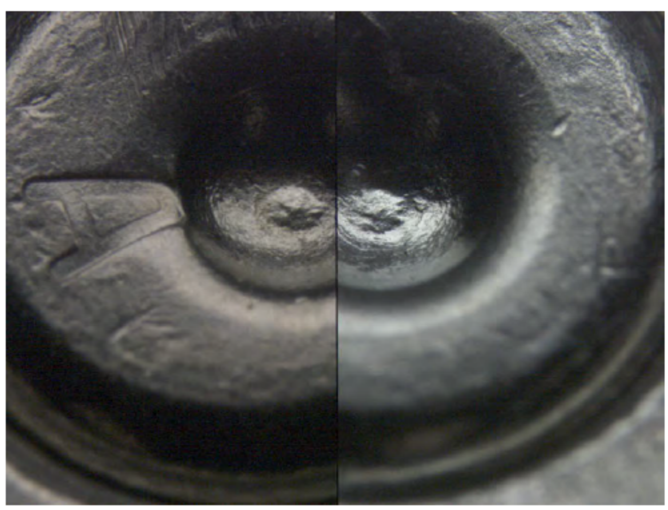


FIG. 6 *Kit #27 Ex 1341 to Ex 1760 28X, FPIM.*

[2] Breechface Mark
[3] Firing Pin Impression Mark

TABLE 3 *Compiled overall study report.*

Total # Kits Distributed	Total # of completed kits returned*	% Participation	# of Laboratories Represented[‡]	Avg. years of Experience	Min Years of Experience	Max Years of Experience
47	34[†]	31/47 0.659 65.9%	22	12.1 years	3 years	46 years

Specimen Population	Total # Identifications Reported	Total # of True Identifications of Kits returned	Total # False Identifications	Total # Comparisons	Total # Inc Reported
Cartridge Cases	191	199	1	693	39
Bullets	156	207	0	955	165

Specimen Population	Total # Eliminations Reported	Total # of True Eliminations of Kits returned	Total # True Eliminations Adjusted	Total # False Eliminations
Cartridge Cases	406	400	441	3
Bullets	519	360	619	1

Specimen Population	Sensitivity	Specificity	Error Rate: False Identification	Error Rate: False Elimination	Overall Error Rate
Cartridge Cases	190/199 0.955	403/441 0.914	1/693 0.144%	3/693 0.433%	5/1648 0.303%
Bullets	156/207 0.754	518/619 0.837	0	1/955 0.105%	

Overall Sensitivity	Overall Specificity
346/406 85.2%	921/1060 86.8%

*This refers to answers that have been submitted not necessarily physical kit.
[†]The data from three kits were not used in the calculations for noted reasons (see report notes page).
[‡]Of returned kits.

Aggregate Analysis

Table 3 summarizes the analysis of data. In addition to the overall error rate, we also measured sensitivity and specificity. Sensitivity was defined as the number of positive conclusions (identifications) actually obtained from the test divided by the number of true positives (true identifications). The sensitivity of a study is important because it relates to the test's ability to identify positive results in this case positive associations of like origin when they exist. It measures the proportion of actual positives that are correctly identified. Specificity was also mea sured in this study. The specificity is the number of negative conclusions (eliminations) actually obtained from a test divided by the number of true negatives possible (true eliminations). The specificity measures the proportion of negatives which are cor rectly identified. This relates to a test taker's ability to properly identify negative results.

Table 3 shows the sensitivity and specificity of the cartridge cases and bullets. The sensitivity and specificity for cartridge cases was 95.5% and 91.4%; for bullets, 75.4% and 83.7%. The false positive and false negative error rate for cartridge case eval uation was calculated by taking the number of false identifications or false eliminations over the total number of cartridge case com parisons made using the most conservative approach. A false positive result is one in which an association is made which is incorrect. Likewise, a false negative result is when an association is not made, when it should be. The false positive error rate recorded for the evaluation of cartridge cases in this study was 0.144%, and the false negative error rate was 0.433%. For bullets, the false positive error rate was 0.0% and false negative error rate was 0.105%. The overall error rate was 0.303%, overall sensitiv ity 85.2%, and overall specificity 86.8% (see Table 4).

A total of 204 inconclusive results (neither identification nor elimination) were reported for the evaluation of the cartridge cases and bullets/bullet jackets in this study, for which a true identification or a true elimination could have been made. Such a response is scientifically valid and acceptable, indicating an insufficient agreement or disagreement of individual microscopic marks of value. It was observed that there were 68 inconclusive responses that should have been identifications and 136 incon clusive responses that should have been eliminations. Of the 68 inconclusive responses that should have been identifications, 62 (91.2%) were for bullets and six (8.8%) for cartridge cases. Of the 136 inconclusive responses that should have been elimina

TABLE 4 *Descriptive statistics for years of experience, sensitivity, and specificity (*N *31).*

		Cartridge		Bullet		Overall	
	Years Exp.	Sensitivity	Specificity	Sensitivity	Specificity	Sensitivity	Specificity
Mean	12.19	0.96 (0.08)	0.93 (0.14)	0.75 (0.23)	0.85 (0.09)	0.85 (0.14)	0.88 (0.10)
Min	3.00	0.71	0.50	0.17	0.67	0.50	0.64
Max	46.00	1.00	1.00	1.00	1.00	1.00	1.00
25%	5.00	0.86	0.93	0.57	0.79	0.67	0.82
50%	11.00	1.00	1.00	0.83	0.80	0.92	0.90
75%	16.00	1.00	1.00	1.00	0.93	1.00	0.96

Standard deviations are shown in parentheses. 95% Confidence Interval for bullet sensitivity 0.68 0.84. 95% Confidence Interval for bullet speci ficity 0.81 0.88. Confidence intervals not reported for cartridge cases due to non normal distributions.

tion responses, 103 (75.7%) were for the evaluation of bullets and 33 (24.3%) for cartridge cases.

There are several variables that can affect how a particular tool marks an object. In this study, these variables include pressure differences between test fires, wear in the microscopic marks, differences in cartridge materials, and use, abuse, and debris, which can create a level of ambiguity in the individual microscopic marks from consecutive test fires within a single firearm. Also, there are internal variables such as the policies and procedures that laboratories use which dictate when an examiner can declare an elimination when the firearm is absent.

Additional Analysis

In this section, we engage in further analysis of sensitivity and specificity, this time looking at results across the 31 examiners. No analysis of error rates is appropriate in this fashion, as there were so few errors that no meaningful variability exists across examiner kits. The questions we address here are as follows: what does the variation in sensitivity and specificity look like across the examiners; and to what extent is there a relationship between years of experience of the examiners and their sensitivity and specificity levels.

Table 4 reports descriptive information on these variables, including their mean, standard deviation, minimum and maximum, and quartiles. Note that for sensitivity and specificity levels, the means will be slightly different than the overall levels reported in the first section of our findings. This is because the kits varied in their denominators, and thus, averaging 31 kits with different denominators will result in means that vary from overall levels that are calculated without taking into account differences across examiners/kits.

Distributions of sensitivity and specificity for bullets and cartridge cases were different. For bullets, both measures were on a normal distribution, according to a one sample Kolmogorov Smirnov test. For cartridge cases, however, the null hypothesis of normality was rejected for both measures. This is primarily because more examiners were likely to have a perfect (1.0) sensitivity and specificity for cartridge cases than for bullets. For bullets, only nine of 31 kits were associated with a perfect sensitivity, whereas for cartridge cases, 23 kits were perfect in sensitivity. Similarly, for specificity, 5 of 31 kits were perfect for bullets, compared to 20 for cartridge cases. Thus, we could only calculate confidence intervals across examiners for bullets, not for cartridge cases. A 95% confidence interval for bullet sensitivity levels ranges from 0.68 to 0.84. A 95% confidence interval for bullet specificity levels ranges from 0.81 to 0.88. Sensitivity ratings, then, varied much more dramatically across examines than did specificity.

Years of experience varied from 3 to 46. Although the sample size was too small to make general conclusions about the relationship between years of experience (YOE) and sensitivity and specificity, we still performed some limited analysis. The correlation between YOE and both sensitivity levels was near zero. However, the correlation between YOE and both specificity levels was approximately 0.25, with a *p* level of 0.08 (not significant at 0.05, but close). To see whether there may be a more complex (rather than linear) relationship between YOE and sensitivity and specificity, we broke the levels down by four categories of YOE, which are consistent with the quartiles in Table 4. Table 5 reports this analysis (for bullets only; no meaningful patterns emerge with cartridge cases). For sensitivity, levels jump up markedly from those at the beginning of their career (0.63 for 1 5 YOE to 0.84 for those with 6 11 YOE), and then tails off back to 0.71 for those with more than 17 YOE. The pattern is quite different for specificity, with a general gradual increase from an average of 0.82 for those with 1 5 YOE to a 0.90 for those with 17+ YOE, with a little movement in the middle categories.

TABLE 5 *Sensitivity and specificity by years of experience.*

Years of Experience	Bullet Sensitivity	Bullet Specificity
1 5 (*N* 8)	0.63	0.82
6 11 (*N* 9)	0.84	0.85
12 16 (*N* 7)	0.81	0.82
17 46 (*N* 7)	0.71	0.90

Discussion and Conclusions

The number of true eliminations and true identifications varied from test to test. This design provided a realistic study approximating how examiners perform their actual case work. In Giroux's study of consecutively manufactured screwdrivers, he took 80 questioned tool marks and eight known tool marks which were produced using three consecutively manufactured screwdrivers (18). Ten questioned tool marks were randomly numbered, and the eight known test marks were sent to eight different examiners. Examiners were asked to compare the known mark to the unknown marks and render a conclusion. Within this test, there were 29 true identifications and 51 true eliminations. The false positive error rate was 0% and false negative error rate 3.4%. The sensitivity was reported as 75.9% and specificity 15.7%, suggesting that examiners are far less likely to eliminate based on the individual characteristics than to make identification when the latter is possible. However, the decision of inconclusive (or no conclusions) is not accounted for in this test. Such a result is common because in the way that this test is constructed a response of no conclusion does not have a direct impact on how the results are tabulated. Such a limitation to the test can produce results that are unrealistic to the nature of typical firearm/tool mark examinations, the prediction of error rate, and the number of actual comparisons made.

In this study, by contrast, it was observed in the tabulation of the results that a cause and effect exists within the scope of the examination when an inconclusive (neither identification nor elimination) response is reported. During the evaluation of the data, it was observed that for the examination of cartridge cases in this study, which is similar to casework, of the 400 true eliminations that existed (within the 31 tests) as the test was originally designed a total of 406 were reported, three of which were false, however, that leaves three above what was theoretically possible. Yet, when the six inconclusive responses that should have been identifications are factored in, there is an adjustment of 41 additional elimination responses that are now possible based on the inconclusive response. As an examiner renders an opinion of inconclusive, they are now obligated to compare additional items within a group that otherwise would not necessarily need to be compared if an identification or elimination had been made. By default, this creates an independent group in the process requiring its own set of comparisons. This is the case in a number of the comparisons made within this study.

For example: Group A consists of items 1, 2, 3, 4; and Group B consists of items 5, 6, 7, 8. Traditionally, within group A, there are three comparisons, and within group B, there are three comparisons and one comparison between groups A and B.

Because the elimination of any one item in Group A to any one item in Group B separates the two groups, therefore theoretically there is only one true elimination possible as designed. However, the reality is that it is possible and also a correct response to evaluate group A and be inconclusive in items 1 and 2 to 3 and 4. Group A (1, 2) and Group C (3, 4) and Group B (5, 6, 7, 8), and now Group A and B eliminate and Group C and B elimi nate, while Group A and Group C are inconclusive. So although as designed, there was only one elimination possible, based on the response of neither identification nor elimination, which is not incorrect, there are now two true eliminations possible. This same cause and effect occurs for each time an inconclusive response exists that should have been an identification.

A re evaluation of the data, taking this information into con sideration, is what produced a higher aggregate specificity mea surement of 91.4% for cartridge cases and 83.7% for bullets, than what would be typically expected, based on previous stud ies. There were 406 cartridge case eliminations reported, three of which were false, leaving 403 reported eliminations (406 3 = 403). There were 41 eliminations created from the six inconclusive responses that should have been identifications, leaving 362 actual elimination responses reported (403 41) ignoring the inconclusive responses. This creates a specificity measurement of 90.5% (362/400). In the case of the bullet eval uation, there were 519 bullet eliminations, one of which was false, leaving 518 reported eliminations. There were 259 elimina tions created from the 62 inconclusive responses that should have been identifications, leaving 259 actual eliminations responses reported (518 259), ignoring the inconclusive responses. This creates a specificity measurement of 71.9% (259/360) for the evaluation of bullets. Such a measurement is consistent with what is expected based on past studies; however, it is not an accurate assessment of the level of comparisons actu ally made in casework. The realistic evaluation shows that as comparisons are made, an examiner becomes more and more specific in his assessment of the information. Although it has been argued that examiners are less likely overall to make elimi nations, the results of this study indicate that in actual casework, overall they are 86.8% likely to make elimination when elimina tion can be made, and that examiners are 85.2% likely to make identification when identification can be made.

While the error rate is the most important measure of the qual ity for forensic comparison examinations, sensitivity and speci ficity are also indicators of a test's quality and should be given fair consideration. The overall results from this study were dif ferent from previous study results. They provide a more accurate data point indicative of the capabilities of the discipline of fire arm and tool mark identification to make conclusive identifica tions and exclusions with regard to the origin of a mark. This study assessed the overall scientific validity and quality of the examination of ammunition components. Although definitely useful in court and of value scientifically, caution should be used when applying these results to estimate error rates in a general ized sense. A number of factors, such as a laboratory's quality assurance program (which includes verifications and peer review), would influence error rates in casework.

The participant pool for this study ($N = 31$) was fairly impres sive when considering how much time and effort each examiner volunteered to the study. A number of the participants had some type of formal CMS training, although pattern matching was pri marily used within the test, with only two participants noting the use of CMS during their examination. With such variability in mind, it only adds weight to the results that indicate that firearm and tool mark identification does follow valid methodology and that proper training provides each examiner with the skills neces sary to make the correct associations.

An official questionnaire responded to by examiners once the test was completed indicated that the general feeling was that this test did take considerably longer to complete than other tests they had taken or had anticipated. The level of difficulty of the test was also commented on as being more difficult than other studies and more representative of actual casework type distribu tion, which was the goal of this study.

During the past several years, significant research has been published in the evaluation of fired ammunition components. This research has included the test fire of firearms numerous times to evaluate the changes in microscopic characteristics observed on the fired bullets and cartridge cases, as well as the test firing of consecutively rifled firearms to determine whether the projectiles could be identified to the barrel from which they were fired. It has been found in every research project involving such examinations that a properly trained firearm and tool mark examiner has the ability to identify a surface marked by a tool back to the particular tool that made the mark, and likewise eliminate a particular tool on the same basis. However, many of these studies have not included the impact of the inconclusive response when evaluating their data. As indicated through this study, a conclusion of neither identification nor elimination adds weight and value to the clear response of identification or elimi nation. Examiners are trained to be more conservative when making their evaluations and a response of inconclusive means that a particular examiner has not seen enough information to say that two items have been marked by the same tool or that they have not been marked by the same tool. Courts should be more inclined to take validation studies into greater consideration when evaluating the probative value of testimony and evidence, when the studies are conducted in a fashion that resembles actual casework.

Acknowledgments

The authors would like to take this opportunity to express their warmest and sincerest thanks to the San Francisco Police Department Criminalistics Laboratory Firearm and Tool Mark Unit for the resources, time, and personnel to complete this research project. An expression of thanks to Ron Nichols for his assistance in the compilation and interpretation of the data acquired during this study. Thanks to Kelsey Ellison, SFPD Fire arm Unit intern, for her assistance in constructing the "labora tory made" bullet retrieval trap, firing of test specimens, and creating test packets. Special thanks to the National Firearms Examiner Academy (NFEA) for the foundational training and work on this project. Additionally, special thanks go out to all of the firearm and tool mark examiners who participated in this study. Participating in this study took a great deal of time and commitment to complete for the advancement of the science.

References

1. Matty W. A comparison of three individual barrels produced from one button rifled barrel blank. AFTE J 1985;17(3):64 9.
2. Brundage D. The identification of consecutively rifled gun barrels. AFTE J 1998;30(3):438 44.
3. Thompson E, Wyant R. Knife identification project. AFTE J 2003;35 (4):366 70.
4. Bachrach B, Jain A, Jung S, Koons RD. A statistical validation of the individuality and repeatability of striated tool marks: screwdrivers and tongue and groove pliers. J Forensic Sci 2010;55(2):348 57.

5. Weller T, Zheng A, Thompson R, Tulleners F. Confocal microscopy analysis of breech face marks of fired cartridge cases from 10 consecu tively manufactured pistol slides. J Forensic Sci 2012;57(4):912 7.
6. Biasotti AA. A statistical study of the individual characteristics of fired bullets. J Forensic Sci 1959;4(1):34 50.
7. Biasotti AA, Murdock JE. Criteria for the identification or state of the art of firearms and toolmark identification. AFTE J 1984;16(4):16 24.
8. Faden D, Kidd J, Craft J, Chumbley LS, Morris M, Genalo L, et al. Sta tistical confirmation of empirical observations concerning toolmark striae. AFTE J 2007;39(3):211 20.
9. Chumbley LS, Morris MD, Kreiser MJ, Fisher C, Craft J, Genalo LJ, et al. Validation of toolmark comparisons obtained using a quantitative comparative, statistical algorithm. J Forensic Sci 2010;55(4):953 61.
10. Neel M, Wells M. A comprehensive statistical analysis of striated tool mark examinations part 1: comparison known matches and known non matches. AFTE J 2007;39(3):176 98.
11. Weavers G, Neel M, Buckleton J. A comprehensive statistical analysis of striated tool mark examinations part 2: comparing known matches and non known matches using likelihood ratios. AFTE J 2011;43(2):137 45.
12. Buckleton J, Nichols R, Triggs C, Weavers G. An exploratory Bayesian model for firearms and tool mark interpretation. AFTE J Fall 2005;37 (4):352 61.
13. Gutowski S. Error rates in the identification sciences. Forensic Bull, 2005;23 9. ISSN 1447 6673.
14. National Academy of Sciences. Strengthening forensic science in the United States: a path forward. Washington, DC: The National Academies Press, 2009.
15. Budlowe B, Bottrell M, Bunch S, Fram R, Harrison D, Meagher S, et al. A Perspective on errors, bias, and interpretation in forensic sciences and direction for continuing advancement. J Forensic Sci 2009;54(4):798 809.
16. Smith ED. Cartridge case and bullet comparison validation study with firearms submitted in casework. AFTE J 2005;37(2):130 5.
17. http://www.random.org/sequences
18. Giroux B. Empirical and validation study: consecutively manufactured screwdrivers. AFTE J 2009;41(2):153 155.

Additional information and reprint requests:
Tasha P. Smith, B.S.
San Francisco Police Department
Criminalistics Laboratory
Firearm and Tool Mark Unit
1245 3rd Street, Building 606
San Francisco, CA 94158
E mail: tbaham06@gmail.com

Supporting Information

Additional Supporting Information may be found in the online version of this article:

Appendix S1. Appendix A.