2015

Estimation of changes in breech face and firing pin marks over consecutive discharges and its impact on 2D correlation systems

Justin Kirk

Follow this and additional works at: <https://researchrepository.wvu.edu/etd>

Recommended Citation

Kirk, Justin, "Estimation of changes in breech face and firing pin marks over consecutive discharges and its impact on 2D correlation systems" (2015). *Graduate Theses, Dissertations, and Problem Reports*. 5985.

<https://researchrepository.wvu.edu/etd/5985>

This Thesis is protected by copyright and/or related rights. It has been brought to you by the The Research Repository @ WVU with permission from the rights-holder(s). You are free to use this Thesis in any way that is permitted by the copyright and related rights legislation that applies to your use. For other uses you must obtain permission from the rights-holder(s) directly, unless additional rights are indicated by a Creative Commons license in the record and/ or on the work itself. This Thesis has been accepted for inclusion in WVU Graduate Theses, Dissertations, and Problem Reports collection by an authorized administrator of The Research Repository @ WVU. For more information, please contact researchrepository@mail.wvu.edu.

**Estimation of changes in breech face and firing pin marks over
consecutive discharges and its impact on 2D correlation systems.**

Justin Kirk

**Thesis submitted
to the Eberly College of Arts and Sciences
at West Virginia University**

in partial fulfillment of the requirements for the degree of

**Master of Science in
Forensic and Investigative Science**

**Keith Morris, Ph.D., Chair
Jacqueline Speir, Ph.D.
Kent Gardner**

Forensic and Investigative Science Program

**Morgantown, West Virginia
2015**

**Keywords: Forensic Science, Firearms, Cartridge Cases, IBIS®
Copyright 2015 Justin Kirk**

ABSTRACT

Estimation of changes in breech face and firing pin marks over consecutive discharges and its impact on 2D correlation systems.

Justin Kirk

When a firearm is discharged, the individual marks of that firearm's breech face and firing pin are imprinted onto the head of the cartridge case. The pattern of these individual marks is reproducible between test fires; however, over a large number of consecutive test fires, it has been observed that minute changes in these individual marks occur. These changes in individual marks have been shown to not affect an examiner in their ability to find an Identification or an Elimination, however, the effect that these changes have on Integrated Ballistic Identification System (IBIS®) is largely unknown. If these changes in individual marks negatively affect IBIS®'s performance, the reliability of IBIS® could be called into question. In order to examine the effect on IBIS®, 200 consecutive test fires were collected from twenty five SR9 Ruger 9mm handguns for a total of 5000 cartridge cases. The cartridge cases were then entered into IBIS® and the full correlation list was recorded for each cartridge case. Evaluation of Receiver Operator Characteristic Curves and Area Under the Curve indicated that the changes in individual marks over a number of consecutive test fires does have a negative effect on IBIS®'s performance. Despite this negative effect, IBIS® still performs well enough to be effective as a screening tool. Secondary objectives of this research include: evaluate IBIS®'s ability to distinguish between Known Match and Known Non-Match, evaluate IBIS®'s ability to distinguish between Delta 190 Plus Known Match and Known Non-Match, determine if the changes from consecutive test fires follow a pattern, determine if larger rate of change is present in new firearms, and evaluate IBIS®'s ability to distinguish between Known Non-Match comparison between two Ruger SR9 handguns and Known Non-Match comparisons resulting from non-Ruger SR9 comparisons.

Acknowledgements

Research was sponsored in part by the U.S. Army Research Office and U.S. Army Research Laboratory and was accomplished under Cooperative Agreement Number W911NF-12-2-0056. 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 Army Research Office, Army Research Laboratory, or the U.S. Government. The U.S. Government is authorized to reproduce and distribute reprints for Government purposes notwithstanding any copyright notation hereon. This project is funded by the Assistant Secretary of Defense for Research and Engineering ((ASD(R&E))) and the project is managed by the Defense Forensics and Biometrics Agency, Office of the Chief Scientist. I would also like to thank all of my teachers at West Virginia University and especially my committee members: Dr. Keith Morris, Dr. Jacqueline Speir, and Kent Gardner. I would also like to thank Mike Bell as well as Dr. Morris's research group for their help in the data collection for this research. Finally, I would like to thank my family and friends who have been very supportive through my graduate schooling.

-Justin Kirk

Table of Contents

1. Introduction	1
1.1 Summary of Experiment	1
1.2 General Firearm Analysis	1
1.3 IBIS® Background	3
1.4 2D Image Correlation Algorithms	4
2. Materials and Methods.....	4
2.1 Test Fire Collection.....	4
2.2 IBIS® Instrumentation Specifics.....	5
2.3 Cartridge Case entry into IBIS®/Data Collection	5
2.4 Image Capture Assembly	6
2.5 Cartridge Case Photography	7
2.6 Image Processing.....	7
2.7 2D Image Comparison.....	8
2.8 Data Organization	8
2.9 Statistical Analysis	9
3. Results	11
3.1 Known Match vs. Known Non-Match	11
3.2 Delta Groups vs. Known Non-Match Scores	12
3.3 IBIS® Score Trends Over Time	18
3.4 Ruger Known Non-Match vs. Other Known Non-Match.....	20
4. Discussion.....	22
4.1 Known Match vs. Known Non-Match	22
4.2 Delta 199 Known Match vs Known Non-Match.....	24
4.3 Effect of Consecutive Test Fires on IBIS® Scores.....	25
4.4 Pattern of Changes Resulting from Consecutive Test Fires	28
4.5 New Firearm’s Effect on Breech Face Changes due to Consecutive Test Fires	29
4.6 Worst-Case Scenario Ruger Non-Match vs non-Ruger Known Non-Match.....	29
4.7 ROC Curve Comparison: IBIS® vs. 2d Image Comparison	31
5. Conclusions	31
6. References	33

List of Figures

FIGURE 1: EXAMPLE OF BREECH FACE IMAGE BEFORE AND AFTER IMAGE PROCESSING	8
FIGURE 2: KNOWN MATCH VS. KNOWN NON-MATCH PROBABILITY DENSITY FUNCTIONS OF VARIOUS IBIS® CORRELATIONS. A) BREECH FACE SCORE; B) FIRING PIN SCORE; C) BREECH FACE RANK; D) FIRING PIN RANK	11
FIGURE 3: KNOWN NON-MATCH VS VARIOUS DELTA KNOWN MATCH GROUPS OF IBIS® SCORES. A) BREECH FACE SCORE; B) FIRING PIN SCORE; C) BREECH FACE RANK; D) FIRING PIN RANK.....	14
FIGURE 4: RECEIVER OPERATOR CHARACTERISTIC CURVES AT VARIOUS GROUPS OF DELTA PLUS KNOWN MATCH GROUPS. A) BREECH FACE SCORE; B) FIRING PIN SCORE; C) BREECH FACE RANK; D) FIRING PIN RANK.....	15
FIGURE 5: AREA UNDER THE CURVE EVALUATION OF RECEIVER OPERATOR CHARACTERISTIC CURVES AT EACH INDIVIDUAL DELTA. A) BREECH FACE SCORE; B) FIRING PIN SCORE; C) BREECH FACE RANK; D) FIRING PIN RANK	18
FIGURE 6: STEPWISE DELTA 1 OF GUN 18 FOR VARIOUS IBIS® CORRELATIONS. A) BREECH FACE SCORE; B) FIRING PIN SCORE; C) BREECH FACE RANK; D) FIRING PIN RANK	19
FIGURE 7: AVERAGE SCORE OF THE STEPWISE DELTA 1 OF ALL FIREARMS FOR VARIOUS IBIS® SCORES. A) BREECH FACE SCORE; B) FIRING PIN SCORE; C) BREECH FACE RANK; D) FIRING PIN RANK	20
FIGURE 8: RUGER KNOWN NON-MATCH VS. OTHER NON-RUGER KNOWN NON-MATCH AT VARIOUS IBIS® SCORES. A) BREECH FACE SCORE; B) FIRING PIN SCORE; C) BREECH FACE RANK; D) FIRING PIN RANK.....	21
FIGURE 9: KNOWN MATCH PROBABILITY DENSITY FUNCTION VS. RUGER KNOWN NON-MATCH PROBABILITY DENSITY FUNCTIONS AT VARIOUS IBIS® SCORES. A) BREECH FACE SCORE; B) FIRING PIN SCORE; C) BREECH FACE RANK; D) FIRING PIN RANK	22
FIGURE 10: RUGER KNOWN NON-MATCH VS OTHER KNOWN NON-MATCH VS. OVERALL DENSITY FUNCTIONS. A) KNOWN MATCH; B) KNOWN NON-MATCH	30

List of Tables

TABLE 1: AVERAGE BAYESIAN LIKELIHOOD RATIOS AT VARIOUS SCORES OR RANKS.	12
TABLE 2: AREA UNDER THE CURVE SCORES OF THE RECEIVER OPERATOR CHARACTERISTIC CURVES IN FIGURE 4	16

1. Introduction

1.1 Summary of Experiment

This research studied how the changes in breech face marks over a number of consecutive discharges affects automated firearm identification systems such as the Integrated Ballistic Identification System (IBIS®). In order to do this, several analyses were performed. To facilitate the experiments, 200 cartridges were consecutively fired from twenty five SR9 Ruger 9mm handguns. The casings were then entered into IBIS® and a pairwise comparison was performed between each of the consecutively fired cartridges. The full IBIS® comparison list was collected for each cartridge to provide a large pool of Known Match and Known Non-Match Scores. For the purpose of this study Known Match and Known Non-Match refers to same gun comparisons and different gun comparisons, respectively. Due to the comparison algorithm of IBIS® being unknown, an additional examination was attempted using a known algorithm. Digital macro photographs of the primer were taken of each cartridge case and a 2D image cross correlation coefficient algorithm was then be used to obtain match scores. The same pairwise comparison will be performed so that the two sets of results can be compared using Receiver Operator Characterizes (ROC) curves. The results were then analyzed to see if the change in breech face and firing pin marks over time has a significant effect on 2D correlation systems. The main objective of this research was to examine if the changes resulting from a number of consecutive discharges affect IBIS® Correlation Scores. The secondary objectives of this research include: evaluate IBIS®'s ability to distinguish between Known Match and Known Non-Match, evaluate IBIS®'s ability to distinguish between Delta 190 Plus Known Match and Known Non-Match, determine if the changes from consecutive test fires follow a pattern, determine if larger rate of change is present in new firearms, determine if Known Non-Match comparison between Ruger SR9 handguns were distinguishable from Known Non-Match comparisons resulting from non-Ruger SR9 comparisons, and compare IBIS® and the 2D correlation using Receiver Operator Characteristic (ROC) Curves.

1.2 General Firearm Analysis

Every time a firearm is discharged, the same mechanical process is followed to fire the bullet and eject the cartridge case. Because of this process, a number of markings are left on the cartridge case that can be used for comparison. The three markings that are commonly used for comparison are the

breech face marks, the firing pin marks, and the ejector marks. This research will focus on the changes on the breech face and firing pin resulting from a number of consecutive test fires.

When a trigger is pulled, the firing pin of the firearm strikes the primer of the cartridge case. The impact of this strike causes the primer to ignite which, in turn, lights the powder in the cartridge case. The burning powder produces gases, thus creating pressure. As the cartridge is in the barrel of the firearm, there are only two directions in which this pressure can travel. As a result, the bullet is projected down and out of the barrel, “firing” the bullet. However, this same force travels in the opposite direction of the bullet, forcing the cartridge into contact with the breech face of the firearm. During this contact, the striations of the breech face are imprinted onto the primer of the cartridge case. Since the cartridge case strikes the breech face with approximately the same force each time the firearm is discharged, the breech face marks are reproducible [7].

While the breech face marks are reproducible, minute changes in the breech face can occur due to wear after a large number of discharges. Additionally, it has been observed that the breech face marks of new firearms can change at a faster rate for a number of rounds that are first fired. This is most likely due to the fact that some of the striations that are present at the firearm’s creation are smaller and more delicate than others. These weaker striations would be more susceptible to wear and thus would change much more rapidly.

This brings us to the central question of this research; can the normal use of a firearm cause a significant enough change in breech face marks or the firing pin impression to affect IBIS®’s ability to correctly classify a Known Match as a potential match? A number of articles exist which state that the change in breech face marks is not significant enough to have an effect on examinations. In a study by Robert Shem (3), 501 consecutive rounds were fired from a .25 ACP Raven pistol. He found that enough individual characteristics persisted on the breech face to reach an identification [3]. In a study by Yoshimitsu Ogihara (4), a similar experiment was performed that involved 5000 consecutively fired cartridge cases from a 45 caliber M1911A1 pistol. He found that the first cartridge case could be positively identified to all of the other 4999 cartridge cases that were fired from that firearm [4]. Finally, a study by Aylin Saribey (6) performed a similar experiment using a number of firearms, firing between 1000 to 14000 consecutive rounds from each firearm. This study also found that the changes in breech face marks are not significant enough to cause false exclusions even after 14000 consecutive fires [6]. While a sufficient of research exists that indicates the change in breech face marks will not cause a false exclusion, very little research exists that examines how these changes affect automated systems such as IBIS®. In an article by Melanie Bernard (5), a study was performed that attempted to find if IBIS®

Correlation Scores are dependent on a number of variables including the number of rounds fired between correlations. They found that IBIS® Correlation Scores were not affected by the difference of up to 25 shots fired between correlation scores [5]. While this research gives a good starting point, it has a small sample size and possessed a number of variables that should be eliminated in a more focused study. This research will be focused on finding if, or to what degree, IBIS® and other correlation algorithms are affected by the changes in individual marks caused by consecutive test fires.

1.3 IBIS® Background

The Integrated Bullet Identification System, or IBIS®, is an automated firearm identification system which consists of two different systems: the Data Acquisition Station (DAS) and the Systems Analysis Station (SAS). The Data Acquisition Station possesses two of IBIS®'s predecessors. The first is BulletProof® which is an acquisition platform for bullets. The second is BrassCatcher™ which is an acquisition platform for cartridge cases. When cartridges are entered into the DAS, the first step is to label it using any class characteristics such as caliber, make, or firing pin type to allow for class characteristic filtration. The DAS then captures two-dimensional digital images of the breech face and firing pin. Once this is completed the images are stored and used by the second system which is known as the Systems Analysis Station or SAS. Once a cartridge is entered into the SAS, it will automatically extract all of the characteristics that can be found in the image. These characteristics will then be searched against every cartridge case in the database with matching class characteristics. An independent search and match score is performed for both the breech face and the firing pin. These scores are referred to as the Breech Face Score and the Firing Pin Score. Two additional variables can be created by ranking all of the Firing Pin or Breech Face Scores between the entered cartridge and each database cartridges in order. These measures are called the Breech Face Rank and Firing Pin Rank. For this study, all four types of scores will be examined separately [9].

Research has been performed to find how examiner variation in cartridge case acquisition affects IBIS® match scores [2]. According to this research, manual or automatic ring selection has no significant effect on match scores. Slight rotational variation around the standard 3 o'clock position was also examined in this research and was found to have no significant effect on match scores. The final parameter that was analyzed was lighting intensity. It was found that there was no statistical difference between 20% and 38% intensity indicating that this is the optimal lighting range [2]. Due to this research, the following measures will be taken to minimize examiner variation: manual ring selection

will be used, each cartridge case will be placed at the standard 3 o'clock position, and the light intensity will be set between 20% and 38%.

The algorithms used by the IBIS® system are proprietary information of FTI Inc. As a result, we can only speculate as to how IBIS® works. Since the IBIS® algorithms themselves cannot be analyzed, it is important to have another commonly used comparison method with a known algorithm. This will also give us the opportunity to compare IBIS® to other automated comparison methods.

1.4 2D Image Correlation Algorithms

A common method for automated cartridge case comparison is to use cross-correlation coefficients. Most notably, this is being used by John Song in his research for NIST titled *Establish the National Ballistics Evidence Search Engine (NBESE) Based on 3D Topography Measurements on Correlation Cells* [8]. Cross-correlation uses the correlation theorem to compare two images and reports how closely those two images match. The simplest form of the correlation equation can be seen below.

$$c(x, y) = \frac{1}{MN} \sum_1^M \sum_1^N f(x, y)g(x + s, y + t) \quad [1]$$

This equation finds the correlation between image f and image g by summing the product of each corresponding pixel value across both images. In this equation, the variables x and y refer to the pixel's location assuming that the image was represented as a Cartesian grid. The variables s and t are implemented in order to allow the images to be shifted on the x and y axis in order to find a better correlation. The variables M and N refer to the total number of pixels on the x and y axis respectively. When using cross-correlation, it is often necessary to normalize the images in order to reduce noise from amplitude differences. This is achieved by subtracting the mean pixel value of the image from each individual pixel. The whole summation must also be divided by the standard deviation of the pixel values of each image. Cross-correlation coefficients are being used by a number of investigators in current research [8, 10].

2. Materials and Methods

2.1 Test Fire Collection

This study focuses on the effects that the changes in breech face marks have on the IBIS® system. For this reason, the ammunition type and make/model of firearm will remain constant. This

research implemented 5000 Remington 9mm Luger cartridges and twenty five SR9 Ruger centerfire pistols. Each firearm was designated a letter and assigned 200 cartridges. When designating letters to firearms, the letters I and O were skipped to reduce confusion with the numbers 1 and 0. The twenty fifth firearm was then designated Gun %. Prior to test firing the rounds, each cartridge was numbered and labeled in the order that they would be fired. The boxes in which the ammunition was purchased were also labeled. The labeling format used was (Designation Letter)### with the three number signs representing the assigned cartridge number. For example, B053 would indicate that the cartridge case was the 53rd cartridge fired from firearm B. A mesh box was used to collect the cartridge cases during test fires. This, along with the cartridges being pre-labeled, allowed for full clips of test fires to be performed before having to stop to reload. All cartridge cases were stored in the ammunition boxes in which they were purchased to keep them organized and to protect the breech faces of the cartridges. All test fires were performed on the same day.

2.2 IBIS® Instrumentation Specifics

The IBIS® equipment used for this experiment consisted of a Remote Data Acquisition Station (RDAS) which allows for on-site data entry, image acquisition and remote correlations. The RDAS is designated SVR1_004_MDEMO. This study used the IBIS® Heritage platform which uses the integrated BrassCatcher™ program. Cartridge case images were captured using a Nikon trinocular head stereoscopic microscope (no. 1104015). This microscope uses a laser based auto-focus and halogen ring lighting (model: LHS-HSOC-11). A Sony CCD camera (XC-ST50, no. 105916) coupled to the microscope was used to acquire images of each cartridge case.

Correlations were performed off-site and sent back to the RDAS. The Quantum system was used as a back-up for all correlation lists received. The back-up tapes were changed each weekday due to the large amount of correlations being performed at the time. Calibration is performed automatically each time the RDAS is logged on. The process resets the X/Y axis motors, focus, and zoom objectives to zero and forces the system to relearn the magnification settings for cartridge case acquisition [9].

2.3 Cartridge Case entry into IBIS®/Data Collection

A Case ID entry was created for each cartridge case. The Case IDs used the following format: RN-JK##-RM-UNK-###. RN stands for Ruger. JK## designates which firearm the cartridge case was fired from. RM stands for Remington. UNK stand for an unknown manufacture of the primer. Finally the last three numbers designate the order in which the cartridges were fired. One Exhibit entry was created in

each Case ID. This method was chosen because IBIS® does not automatically generate a score between other exhibits in the same Case ID. Each cartridge case was entered into IBIS® using Manual ring selection, the standard 3 o'clock position, and a light intensity set between 20% and 38%.

Once all of the cartridges had been entered, all of the correlation lists that IBIS® had automatically performed were deleted. This was required due to IBIS® only computing correlation scores between the entered cartridge and any cartridges that were entered before it. In order to delete all the correlation lists at once, IBIS® support was contacted. Once all of the old correlation lists were deleted, a new correlation list was generated for each cartridge. This allowed for the cartridge case to be compared to all of the cartridges in this study regardless of the order that they were entered. The full correlation list was obtained from the IBIS® for each cartridge case entered which consisted of approximately 2000 comparisons for each cartridge. The correlation lists were obtained from the IBIS® as text files containing the Firing Pin Rank, Case ID of the Database Cartridge, Exhibit Number of the Database Cartridge, Breech Face Score and Firing Pin Score for each comparison between the sample cartridge and the cartridges obtained from the database.

For each firearm, the 200 text files for each cartridge were consecutively ordered and then combined into one text file that represents each comparison resulting from that firearm. The text file was then converted to an excel file and identifying information for each comparison was added. This included information such as: Case ID of the Entered Cartridge, Exhibit Number of the Sample Cartridge, Serial Number of both the Sample and Database Firearm, etc. Excel was also used to find the Breech Face Rank based off the collected Breech Face Scores. Once the excel files were completed, they were converted to Excel Comma Separated Value (CSV) to allow for easy use with R Studio.

2.4 Image Capture Assembly

Because it is impossible to obtain the images acquired by the RDAS, an additional image capture assembly was required in order to obtain images of the cartridge cases that could be used for an independent 2D image comparison. The image capture assembly consisted of the following: a Nikon D90 Camera, a Fotodiox Macro Bellows extension, a 50mm lens with manual F-stop adjustment, a Velbon Super Mag Slider vertical rail, and a Bencher Copy Mate II copy stand with two built in light sources. A moving stage was fixed in place that allowed for macroscopic adjustments in 3 dimensions. DIYPhotoBits was used to control the camera remotely in order to eliminate vibrations resulting from pressing the shutter release.

2.5 Cartridge Case Photography

Multiple F-Stops were examined to find the setting that provided the best depth of field while still providing a fast enough shutter speed to eliminate blur resulting from minor vibrations acting on the camera. It was found that F11 provided the best results. The camera was set to aperture priority so that the camera would automatically adjust the shutter speed to provide the correct exposure. Due to the bellows and lens being used, the camera had to be manually focused for each photograph. Each cartridge was oriented in the standard 3 o'clock orientation. Each image obtained was labeled using the format described above. All images were taken in the RAW format and later converted to TIFF files using the ViewNX 2 that was packaged with the camera.

2.6 Image Processing

Once the images were obtained, a number of modifications were made in Photoshop to prepare them for comparison. The first modification was a greyscale filter to eliminate color as a variable in the comparisons. The second modification used a Laplacian of Gaussian filter, seen below, to sharpen the images. This provided more contrast in the striations.

$$\begin{bmatrix} 0 & 0 & -1 & 0 & 0 \\ 0 & -1 & 1 & -1 & 0 \\ -1 & 1 & 5 & 1 & -1 \\ 0 & -1 & 1 & -1 & 0 \\ 0 & 0 & -1 & 0 & 0 \end{bmatrix}$$

Laplacian of Gaussian Filter

Third, the firing pin and the area around the striations located around the firing pin were masked. A mask template was used to ensure the same amount of space was masked in each image. Finally the area around the primer was cropped so that each image was exactly 1500 by 1500 pixels with the masked firing pin area directly in the center of the image. The end result can be seen in *Figure 1* below.

Figure 1: Example of Breech Face Image before and after Image Processing

2.7 2D Image Comparison

Before the breech face images could be compared, a separate program was required to register, or orientate, the images to each other. Due to the fact that no non-proprietary program exists to register images of cartridge case breech faces, a pre-existing registration software coded to compare shoe prints was tested to see if it could be applied to cartridge case images. It was found that this program was not able to properly register the cartridge cases. Due to time restraints, this portion of the research was discontinued in order to focus on the IBIS® section of the study.

2.8 Data Organization

The data was organized into the following categories. The Breech Face Score, Firing Pin Score, Breech Face Rank, and Firing Pin Rank were treated as individual variables and examined separately. For each variable the following data sets were created: a Known Match pool, a Known Non-Match pool, Known Match pools at specific Deltas, Stepwise Known Match set, a Ruger Known Non-Match pool, and non-Ruger Known Non-Match pool. For the purpose of this study Known Match and Known Non-Match refers to same gun comparisons and different gun comparisons, respectively

The Known Match pool simply contains all values of comparisons that are known to be matches. Likewise, the Known Non-Match pool contains all values of comparisons that are known to not be matches.

The Delta data sets involve groups of Known Match values that have a specific number of test fires between the two cartridges being compared. For example, the Delta 2 pool will contain the comparison between the first and the third cartridge fired from that firearm. It will also contain the comparison between the second and fourth, third and fifth, etc. A separate data set was created for Delta 1 through Delta 199. This allows for the analysis of how the number of consecutive test fires effects the IBIS® scores.

The Delta 1 Known Match pairs are also referred to as the Stepwise Known Match subset when the comparisons are being examined in the order they were fired. For each firearm tested, the first point would be the comparison between test fire 1 and test fire 2. The next point would be between 2 and 3 and so forth until the last point comparing 199 and 200.

Finally, the Ruger Known Non-Match pool will contain all comparisons that are known to not be a match specifically between Ruger firearms. The non-Ruger Known Non-Match pool will contain all of the remaining comparisons that are known not to be a match that are between the cartridge and any non-Ruger firearm. This will provide a way to see if the similar class characteristics of the Ruger firearms affect the Known Non-Match Scores of IBIS®.

2.9 Statistical Analysis

One of the main methods of evaluation that has been used in this research is the Receiver Operating Characteristic (ROC) Curve which is often used in the medical field for evaluating the performance of a diagnostic test. A ROC curve is a plot of a system's true positive rate (sensitivity) versus the system's false positive rate (1-specificity) [11]. The true positive rate (TPR) represents the percentage of the time that a positive comparison is reported as positive while the false positive rate (FPR) represents the percentage of time that a negative comparison is reported as negative. In any diagnostic test, a threshold can be set in order to find a fixed TPR and FPR. This threshold can be increased or decreased in order to improve either the TPR or FPR; however, improving one will hurt the other. ROC curves plot TPR versus FPR obtained using every available threshold. This provides a good visual representation of the system's performance [11].

In order to provide a quantitative representation of a system's performance, the Area Under the Curve (AUC) can be calculated from a ROC curve. The AUC is the area between the ROC curve and the x-axis. This value can range between 0.5 and 1. It has been shown that the AUC indicates the probability of a system correctly ranking a Known Match comparison above a Known Non-Match comparison. This indicates that the AUC is similar to the Wilcoxon or Mann-Whitney statistical test [12]. In order to find if

two AUC values are statistically different, the Standard Error must first be calculated for each system using the following equations.

$$SE(AUC) = \sqrt{\frac{AUC(1 - AUC) + (n_A - 1)(Q_1 - AUC^2) + (n_N - 1)(Q_2 - AUC^2)}{n_A * n_N}}$$

$$Q_1 = \frac{AUC}{(2 - AUC)}$$

$$Q_2 = \frac{2 * AUC^2}{(1 + AUC)}$$

[12]

The variable n_A represents the sample size of Known Matches being analyzed and n_N represent the sample size of Known Non-Matches being analyzed. Once the Standard Error has been calculated, a Z value can be calculated between the two systems using the following equation.

$$z = \frac{|AUC_1 - AUC_2|}{\sqrt{SE_1^2 + SE_2^2}}$$

[13]

The Z value is then used to obtain a 2 sided P value which is then evaluated at a significance of 0.05. This indicates whether the two AUC values are significantly different with a 95% confidence.

3. Results

3.1 Known Match vs. Known Non-Match

Figure 2 represents the comparisons between the Known Match probability density function and the Known Non-Match probability density function of the IBIS® Scores. As expected, the Known Match functions can be distinguished from the Known Non-Match functions. However, it is important to note that there is a large overlap. The Breech Face Score exhibits the worst separation in density functions. In order to evaluate what a given score can tell us, we can use Bayes Theorem to find a Average Likelihood Ratio at any given score or rank.

Figure 2: Known Match vs. Known Non-Match probability density functions of various IBIS® Correlations. a) Breech Face Score; b) Firing Pin Score; c) Breech Face Rank; d) Firing Pin Rank

Table 1 represents the Average Bayesian Likelihood Ratios for each variable at various scores. These results show that Breech Face Scores above 181 or Firing Pin Scores above 135 very strongly

supports that the comparison is a true match. The results also show that any Breech Face score below 75 or Firing Pin Score below 65 supports that the comparison is not a match. The results for rank provide less strength when considering a comparison. It should be noted that, while the Breech Face and Firing Pin Rank breaks into the 101-1000 category, the maximum Average Likelihood Ratio for both categories does not exceed 150.

Table 1: Average Bayesian Likelihood Ratios at various scores or ranks.

IBIS® Score Type	Average Likelihood Ratio				
	< 1	1-10	11-100	101-1000	1001+
BF Score	0-75	76-100	101-120	121-180	181+
FP Score	0-65	66-88	89-110	111-134	135+
BF Rank	300+	66-300	6-65	1-5	-
FP Rank	351+	91-350	13-90	1-12	-

3.2 Delta Groups vs. Known Non-Match Scores

Figure 3 represents the probability density functions of the Known Match, Known Non-Match, and groups of Known Match above a given Delta. For example, Delta 100+ represents all Known Match Scores at or above Delta 100. From this figure we can see that, as Delta rises, the probability density function of the remaining Known Match Scores shift closer and closer to the Known Non-Match density function. The increase of Delta affects the Breech Face Score the most. This resulted in small shifts up to Delta 150+ and a large shift in the final Delta 190+ that is very similar to the Known Non-Match probability density function. The Firing Pin Score probability density function shift is less drastic. There are small shifts at each Delta step and the final Delta 190+ probability density function is still easily distinguishable from the Known Non-Match probability density function. Both Rank variables show a drop in in the Known Match probability density functions, however, there is still a separation between the Known Match and Known Non-Match below a rank of 100. While the effect of the consecutive discharges on IBIS® Scores can be seen, this does not give us a measureable way to quantify this change.

a.

b.

c.

d.

Figure 3: Known Non-Match vs various Delta Known Match Groups of IBIS® Scores. a) Breech Face Score; b) Firing Pin Score; c) Breech Face Rank; d) Firing Pin Rank

Figure 4 represents Receiver Operator Characteristic Curves (ROC curves) that have been generated comparing the same groups of Known Match Scores in Figure 3 against the Known Non-Match Scores. Visually, one can see that, as Delta increases, the IBIS® scores ability to distinguish Known Match and Known Non-Match deteriorates. Specifically, it can be seen that IBIS® Firing Pin Score and Rank are not greatly affected by the increase in delta. The Breech Face Score and Rank, however, show a gradual decrease in IBIS®'s ability to distinguish Known Match and Known Non-Match as delta increases. In order to evaluate the overall effectiveness of each ROC curve, the Area under the Curve (AUC) was calculated.

a.

c.

b.

d.

Figure 4: Receiver Operator Characteristic curves at various groups of Delta Plus Known Match Groups. a) Breech Face Score; b) Firing Pin Score; c) Breech Face Rank; d) Firing Pin Rank

Table 2 represents the Area under the Curve (AUC) score obtained from each ROC curve in Figure 4. An AUC score of 1 would represent a system that is never wrong while a score of 0.5 would represent a system that cannot distinguish between Known Match and Known Non-Match at all. From these results we can see that the Breech Face Score and Rank ability to distinguish Known Match and Known Non-Match steadily declines as delta+ increases. Conversely, the Firing Pin Score and Rank show very little decline as delta+ increases. When the AUC scores are compared to each other the following results can be reached. The Breech Face Score and Rank AUC were found to be significantly different between each stepwise comparison with 95% confidence. The Firing Pin Score and Rank AUC scores,

however, were not found to be significantly different between the stepwise comparisons with 95% confidence. A significant difference was found when comparing the Firing Pin Known Match AUC with the Delta 150+ AUC. This gives us an overall scope of how the change in individual marks can affect IBIS® ability to distinguish Known Match and Known Non-Match. For a more in-depth examination, an AUC score must be calculated for each Delta.

Table 2: Area under the Curve Scores of the Receiver Operator Characteristic Curves in Figure 4

IBIS® Score Type	AUC Score			
	KM	Delta 100+	Delta 150+	Delta 190+
BF Score	0.734	0.707	0.684	0.626
FP Score	0.820	0.811	0.804	0.792
BF Rank	0.763	0.732	0.706	0.657
FP Rank	0.851	0.845	0.842	0.825

Figure 5 represents the AUC score resulting from the ROC curve of each individual Delta vs. the Known Non-Match Scores. Figure 5a shows the results of the Breech Face Scores. It can be seen that there is a drop in AUC scores over the first 25 Deltas. This indicates that the changes in breech face marks can actually affect IBIS® after only a few consecutive fires. After Delta 25, the AUC scores level out and drop at a much lower pace. This could indicate that, while IBIS® can see the small changes resulting from cartridge to cartridge, the overall pattern it is extracting from the breech face marks is remaining the same over a large number of consecutive fires. Once we reach Delta 150, the AUC score begins to drop more rapidly again. This indicates that the small changes that affect IBIS® in the earlier Deltas begin to add up and affect the overall pattern that was steadily declining up to this point. When the AUC scores found in Figure 5a are compared to each other the following results can be reached. None of the comparisons between Delta x and Delta x+1 were found to be significantly different with 95% confidence. The comparisons between Delta 1 and Delta 3 through Delta 199 were all found to be significantly different. Because of the initial drop in AUC score over the first 25 Delta, a comparison between Delta 25 and Delta 26 through Delta 199 was also performed. It was found that Delta 25 was not significantly different from Delta 26 through Delta 45; however, comparisons beyond Delta 46 were found to be significantly different. Figure 5b shows the results for the Firing Pin Scores. The Firing Pin Scores do not show the initial drop that we could see in the Breech Face Scores. The Firing Pin Score also starts at a score 0.05 higher than the initial Breech Face Score. As the Delta increases, the AUC score

shows a slow and steady decline. At Delta 175, the AUC scores start to show more variation and start to decrease at a more rapid pace. When the AUC scores found in Figure 5b are compared to each other the following results can be reached. None of the comparisons between Delta x and Delta $x+1$ were found to be significantly different with 95% confidence. The comparisons between Delta 1 and Delta 3 through Delta 199 were also examined. It was found that Delta 1 was not significantly different from Delta 3 through Delta 21; however, comparisons beyond Delta 22 were found to be significantly different. Figure 5c shows the results for the Breech Face Rank. The Breech Face Rank starts at an AUC score 0.05 higher than the Breech Face Score. Other than the initial AUC score, the Breech Face Rank AUC scores follow the same pattern as the Breech Face Score. A sharp drop is present in the first 25 points which then levels off to a small decrease over time. This small decrease continues until Delta 150 where the scores begin to decrease more rapidly and show more variation in score. When the AUC scores found in Figure 5c are compared to each other the following results can be reached. None of the comparisons between Delta x and Delta $x+1$ were found to be significantly different with 95% confidence. The comparisons between Delta 1 and Delta 3 through Delta 199 were all found to be significantly different. Because of the initial drop in AUC score over the first 25 Delta, a comparison between Delta 25 and Delta 26 through Delta 199 was also performed. It was found that Delta 25 was not significantly different from Delta 26 through Delta 50; however, comparisons beyond Delta 51 were found to be significantly different. Figure 5d shows the results for the Firing Pin Rank. The Firing Pin Rank starts at an AUC score 0.05 higher than the Firing Pin Score. Other than the initial AUC score, the Breech Face Rank AUC scores follow the same pattern as the Firing Pin Score. The AUC Score shows a slight and steady decrease from Delta 1 to Delta 175 and then starts to decrease more rapidly and show more variation in score. When the AUC scores found in Figure 5d are compared to each other the following results can be reached. None of the comparisons between Delta x and Delta $x+1$ were found to be significantly different with 95% confidence. The comparisons between Delta 1 and Delta 3 through Delta 199 were also examined. It was found that Delta 1 was not significantly different from Delta 3 through Delta 30; however, comparisons beyond Delta 31 were found to be significantly different. It should be noted that in all four graphs, the AUC scores start to vary to a greater degree after Delta 175. This increased variation could be due to the smaller sample size available as Delta increases. Due to the small sample size, conclusions based on the data represented above Delta 190 should be very conservative in nature.

a.

c.

b.

d.

Figure 5: Area under the Curve Evaluation of Receiver Operator Characteristic curves at each Individual Delta. a) Breech Face Score; b) Firing Pin Score; c) Breech Face Rank; d) Firing Pin Rank

3.3 IBIS® Score Trends Over Time

Figure 6 represents the Stepwise Delta 1 Correlation Scores for Gun 18 for each IBIS® variable. In order to determine if the changes due to a number of consecutive discharges followed a pattern, the Stepwise Delta 1 Known Match comparisons were examined for each firearm separately. Stepwise Delta 1 refers to each Delta 1 comparison for a specific firearm in the order they were fired. For example, the first Stepwise Delta 1 point will represent the comparison between the first and second cartridge fired from that firearm while the fifth Stepwise Delta 1 point will represent the comparison between the fifth and sixth cartridge. A separate graph was created for all four IBIS® Scores for each firearm. From examining these graphs, no discernible pattern could be found. Gun 18 was chosen as a representation because it was the set of graphs that came closest to presenting a pattern.

a.

c.

b.

d.

Figure 6: Stepwise Delta 1 of Gun 18 for various IBIS® Correlations. a) Breech Face Score; b) Firing Pin Score; c) Breech Face Rank; d) Firing Pin Rank

In order to continue searching for a pattern in the changes caused by consecutive discharges, the Stepwise Delta 1 data sets for each firearm were combined. Figure 7 represents the average and 95% confidence interval of at each Stepwise Delta 1. From this data, no discernible pattern in the changes caused by consecutive discharges can be found. From this data, it can be seen that the Firing Pin Score has less variation than the other three IBIS® Scores.

a.

c.

b.

d.

Figure 7: Average score of the Stepwise Delta 1 of all firearms for various IBIS® Scores. a) Breech Face Score; b) Firing Pin Score; c) Breech Face Rank; d) Firing Pin Rank

3.4 Ruger Known Non-Match vs. Other Known Non-Match

Figure 8 represents the probability density functions of Known Non-Match Breech Face Scores produced from other Ruger SR9 handguns and Known Non-Match Breech Face Scores of non-Ruger firearms. From the probability density functions of the Breech Face Score and Breech Face Rank, we can see that IBIS® tends to give higher scores to Ruger Known Non-Match than Known Non-Match from other firearms. This provides evidence that the matching class characteristics of the Ruger SR9 can have an effect on the IBIS® when examining the Breech Face. When examining the Firing Pin Score and Rank, the probability density functions overlap which indicates the matching class characteristic of the Ruger SR9 firing pin does not have an effect on IBIS® Firing Pin Correlations.

a.

c.

b.

d.

Figure 8: Ruger Known Non-Match vs. Other non-Ruger Known Non-Match at various IBIS® Scores. a) Breech Face Score; b) Firing Pin Score; c) Breech Face Rank; d) Firing Pin Rank

Since the probability density functions of Figure 8 did show that it is possible for class characteristics of the Ruger SR9 to have an effect on the IBIS® Breech Face Correlations, the Known Match probability density functions were compared to the Ruger only Known Non-Match to investigate whether these probability density functions could still be distinguished. Figure 9 represents the probability density function of Known Match and Ruger Known Non-Match. We can see from these results that the probability density functions can still be distinguished despite the Ruger Known Non-Match Scores being a worst case scenario. It should be noted that this figure is very similar to Figure 2. This could be due to the large number of Ruger Known Non-Match Scores in the overall Known Non-Match Score pool of data.

a.

b.

c.

d.

Figure 9: Known Match probability density function vs. Ruger Known Non-Match probability density functions at various IBIS® scores. a) Breech Face Score; b) Firing Pin Score; c) Breech Face Rank; d) Firing Pin Rank

4. Discussion

4.1 Known Match vs. Known Non-Match

The first objective of this research was to determine if IBIS® is capable of distinguishing Known Matches from Known Non-Matches. In order to examine this, the Known Match and Known Non-Match probability density functions in Figure 2 were created. From this figure, it can be seen that the probability density functions can be distinguished despite significant overlap in some areas. Of the four variables examined, the Firing Pin Score (Figure 2b) showed the most distinct separation of probability density functions. The probability density functions of the Breech Face Score (Figure 2a) show much more overlap despite being distinguishable from one another. The Breech Face and Firing Pin Rank were also examined. With both Rank variables (Figure 2c and 2d), the Known Match and Known Non-Match probability density functions are more distinguished from each other than the Score variables. This could indicate that some correlation list created by an IBIS® comparison result in lower or higher scores

for every correlation in that list. This could be causing the large overlap in the Breech Face Score probability density functions by providing high Known Non-Match Scores that, while lower than the Known Match Scores in a given correlation list, are higher than Known Match Scores in another correlation list. Due to this, the rank variables are acting as a normalized version of the score variables.

In order to find how reliable a comparison is given a specific score or rank, Average Bayesian Likelihood Ratios (Table 1) were calculated using the two probability density functions from each variable. These Likelihood Ratios show that, for both Breech Face and Firing Pin Scores, there is a point at which a high score can strongly indicate that the comparison is more likely a match than a non-match despite the variation that seem to be present between correlation lists. For Breech Face Score this threshold is 120. The threshold for the Firing Pin Score is 110. The low Likelihood Ratios calculated from the Rank variables indicated that the rank of the comparison alone does not provide a good indication of if the comparison is a match or not.

Another method of evaluating how effectively IBIS® differentiates Known Match vs Known Non-Match is to examine the ROC curves in Figure 3. The ROC curve that was created using the full Known Match pool is represented by the blue line. From these ROC curves we can see that the Firing Pin Correlation still provides a better indicator of whether a comparison is a match than the Breech Face Correlation. In order to provide a way to compare the ROC curves and evaluate the performance of the system, the Area under the Curve (AUC) was calculated for each ROC curve. In this case, the term performance refers to IBIS® ability to correctly allocate a higher score to Known Matches than the scores given to Known Non-Matches. The AUC score for each of the Known Match ROC curves can be found in the KM column of Table 2. From these scores we can see that the Firing Pin Rank provides the best distinction between Known Match and Known Non-Match. We can also see that the rank variables provide better performance than their score counterparts. This indicates that the rank variables are acting as a normalized version of the score variables. Upon investigation, it was found that the range of correlation scores varies between correlation lists. Despite this, the Known Match correlations still tended to be a designated a higher score than the Known Non-Matches correlations. Due to this, combining correlation lists results in a number of low ranking Known Non-Matches with high scores and a number of high ranking Known matches with low scores which distorts the data. Fortunately, the rank variables acting as a normalized version of the score variables eliminate this distortion when performing ROC and AUC evaluations. This could explain why rank performs better in these evaluations and indicates that the rank variables are a more accurate measure of IBIS®'s performance.

Overall, we can determine that IBIS® is capable of differentiating between Known Matches and Known Non-Matches. The Firing Pin Score and Rank seems to provide a better differentiation than that of the Breech Face Score and Rank. IBIS® ability to differentiate is powerful enough that Breech Face or Firing Pin Scores above certain thresholds can strongly indicate that the comparison is a match despite variations in scores between correlation lists. While the Rank variable results provide a good ability to distinguish between Known Match from Known Non-Match, Rank alone does not provide a good indication of whether a comparison is a match or not.

4.2 Delta 199 Known Match vs Known Non-Match

The second objective of this research involves examining IBIS®'s ability to distinguish Known Match vs Known Non-Match under the worst case scenario that the data set could provide. Originally the Delta 199 known match group was going to be compared to the Known Non-Match group, however, it possessed too small of a sample size to be used alone. Instead the Delta groups 190 through 199 were pooled together and designated Delta 190+. To begin, the probability density functions of the Delta 190+ Known Matches (red) was calculated and compared to the Known Match (blue) and the Known Non-Match (black). The results (Figure 3) varied greatly depending on the variable being examined. The Breech Face Score Delta 190+ probability density function shows a significant shift towards the Known Non-Match probability density function. While the two probability density functions can still be distinguished from each other, the amount of shift in the probability density function does show that the change in breech face marks does have a significant effect on the IBIS® Breech Face Correlation. The Firing Pin Score Delta 190+ probability density function showed a much less dramatic shift towards the Known Non-Match probability density function. It can be seen that the Firing Pin Correlation is affected by the changes resulting from consecutive discharges but to a much smaller degree than that of the Breech Face Correlations. The Rank variables show a very similar pattern. With the Breech Face Rank, the spike seen between Rank 1 and 100 is much less prominent after the consecutive test fires, however separation is still present. With the Firing Pin Rank the initial Known Match spike drops but, like with the Firing Pin Score, is not effected as much as the Breech Face Rank. The jagged appearance of the Delta 190+ probability density functions in both rank variables should be noted. This could indicate that the sample size for these Delta groups is not large enough to represent a range of data spanning from 1 to 2000.

In order to further investigate how IBIS® is affected by the changes in breech face, Figure 3 can again be examined. The ROC curves that were created with the Delta 190+ Known Match verses the

Known Non-Match are indicated with a red line. From these ROC curves it can be seen again that the Firing Pin Score is much more resistant to the changes resulting from consecutive test fires. From the AUC scores in Table 2 we can see that both the Breech Face Score and Rank dropped by approximately 0.1. This represents a significant drop in IBIS®'s ability to differentiate between Known Match and Known Non-Matches resulting from the changes that occur over 200 consecutive discharges. The Firing Pin Score and rank only drops by approximately 0.025 which indicates that it is resistant to change over 200 consecutive test fires.

Overall, we can see that the changes occurring from a number of consecutive discharges do have a significant effect on the IBIS® Correlations. The Firing Pin Correlations seem to be effected by the changes in individual marks to a much smaller degree then the Breech Face Correlations. Despite the effect of these changes, it is still possible for IBIS® to find Known Matches out of a group of Known Non-Matches and can still be effective as a screening tool.

4.3 Effect of Consecutive Test Fires on IBIS® Scores

The third and main objective is to perform a more thorough examination of the effect that changes resulting from consecutive test fires has on IBIS®. In the first two objectives, IBIS®'s overall ability to distinguish the full data set was examined as well as the worst case scenario that the data set could provide. While these two situations are important areas to specifically focus on, it is also important to evaluate the data in between to find the point at which IBIS® starts to become significantly affected by the changes resulting from consecutive test fires. To begin, the probability density functions found on Figure 3 should be examined. With the Breech Face Score it can be seen that the shift in Known Match probability density function between the full Known Match and the Delta 100+ Known Match probability density function is small and much less pronounced than the shift seen in the Delta 190+ Known Match probability density function. Similarly, the Delta 150+ Known Matches represent another small shift toward the Known Non-Match probability density function that is a less pronounced shift then that seen in the Delta 190+ Known Match probability density function. The Firing Pin Score shows a similar pattern on a smaller scale. The shifts present from the full Known Match probability density function to Delta 100+ and from Delta 100+ to Delta 150+ are smaller than the shift observed between Delta 150+ to Delta 190+. The Firing Pin Rank probability density functions observed in Figure 3 represents a shift toward the Known Non-Match probability density function between the Known Match probability density function and the Delta 100+ probability density function. The probability density

functions then seem to remain constant from Delta 100+ to Delta 190+. As mentioned in Section 4.2, the jagged probability density functions observed in both Rank variables for Delta 190+ should be noted and could indicate that the sample size is too small to cover the large range of potential values.

Next the ROC curves represented in Figure 4 should be examined. The ROC curves for each variable show a similar trend as the probability density functions. A uniform shift is present between the Known Match and Delta 100+ ROC curves and between the Delta 100+ and Delta 150+ ROC curves. A larger shift can then be seen between the Delta 150+ and the Delta 190+ ROC curves. Further study that would focus on an increase in the number of consecutive test fires, and therefore increasing the sample size at Delta 190+, is recommended. In order to examine the change in Delta's effect on IBIS® performance further, the AUC scores in Table 2 can be examined. The results provide a more quantitative representation of what could be visually seen in Figure 4. With the Breech Face Score, it can be seen that the AUC score drops approximately 0.025 between the full Known Match and Delta 100+ ROC curves and the Delta 100+ and Delta 150+ ROC curves. These shifts in AUC were found to be significantly different with 95% confidence. The amount that the AUC score drops then doubles to approximately 0.06 between the Delta 150+ and the Delta 190+. This represents that up until Delta 150, IBIS®'s performance is being slowly diminished as the Delta grows larger. However, at some point past Delta 150 IBIS®'s performance begins to be affected to a larger degree. The same trend can be seen in the Breech Face Rank. When examining the Firing Pin Score, a different trend is observed. All three shifts observed in the Firing Pin Score AUC scores are approximately 0.01. Furthermore, none of these AUC scores were found to be significantly different with 95% confidence. This indicates that, in regard to the Firing Pin Score, a number of consecutive test fires effect on IBIS®'s performance is both small and relatively constant as Delta increases. However, the comparison between Known Match AUC and Delta 150+ AUC was found to be significantly different with 95% confidence. The AUC scores for the Firing Pin Rank follows the same trend as the Firing Pin Score. It should be noted that both rank variables perform better than their score counterparts. When the Score Known Match AUC was compared to their Rank Known Match AUC counterparts, they were found to be significantly different with 95% confidence. This supports the hypothesis that the rank variables are acting as a normalized version of the score variables. It should also be noted that the Firing Pin Correlations perform better than the Breech Face Correlations.

Due to the results observed in Table 2, it was decided that a more thorough examination was required. In order to do this, an AUC score was calculated for each individual Delta Group. These AUC scores were then ordered from Delta 1 to Delta 199 and represented in Figure 5. The first thing to note

is that the score variables possess the same pattern as their respective rank variables. The only significant difference being that the Rank variables seem to possess a constant increase in AUC score at each Delta over their score counterpart. The Breech Face Score and rank AUC begins with a sharp drop in AUC score between Delta 1 and Delta 5 which indicates that the breech face marks changing to some degree from each test fire and that IBIS® Breech Face Correlations are specific enough to detect these changes. This could represent the variance in individual characteristics being reproduced by the breech face. After the initial drop, however, the AUC scores level out and begin to slowly decrease as Delta increases with this trend continuing until Delta 150. This indicates that the majority of the features that IBIS® is using to calculate the Breech Face Correlation are remaining the same despite the minor differences that are affecting it to some degree. Also as Delta increases the effect of these smaller changes slowly begin to affect the overall pattern that IBIS® is detecting which causes the slow decrease in AUC scores over time. Beyond Delta 150, the AUC score begins to decrease more rapidly and varies to a greater degree. The Firing Pin Score and rank do not exhibit the initial drop in AUC score between Delta 1 and Delta 5 that was observed in the Breech Face variables. Instead it exhibits a slow decrease in AUC score as Delta increase similar to what is observed in the Breech Face variables between Delta 5 and Delta 150. In the Firing Pin variables this trend continues from Delta 1 to Delta 170. Beyond Delta 170, an increase in variance is observed. This indicates that, unlike the Breech Face Correlations, the Firing Pin Correlations are not affected by the changes that occur from a small number of test fires. Instead, it is only affected by these changes building up over a very large number of test fires. For all variables, it should be noted that the higher Delta becomes, the smaller its sample size becomes. This could be the cause of the increased variance in the AUC score and indicates that further research should be performed with more emphasis on examining higher Delta groups.

Overall the results of this research indicate that the changes occurring due to a number of consecutive discharges do affect the performance of IBIS® Correlation Scores. The Breech Face Correlations are affected by these changes to a greater degree than Firing Pin Correlations. Furthermore, IBIS® shows an initial significant drop in performance between the first five discharges before leveling off. Despite this, IBIS® Breech Face Correlations remains accurate enough to justify its use as a screening tool. The Firing Pin Correlation's performance exhibited a much greater resistance to the changes resulting from a number of consecutive discharges. Both correlations' AUC scores exhibited an increase in variance as the higher Delta groups were examined which is likely due to the sample size becoming smaller. Further research should be performed that focuses on providing a larger sample size at larger Deltas.

Furthermore, it should be noted that the method used for collecting test fires could be affecting the results. Since all 200 test fires were performed in quick succession for each firearm, the heating of the firearm could have increased the amount of change occurring in the breech face. In future research, the test fire collections could be broken into multiple days of shooting 50 rounds from each firearm in order to simulate the amount of rounds a typical firearm owner might expend when target shooting.

4.4 Pattern of Changes Resulting from Consecutive Test Fires

The fourth objective of this research is to examine whether the changes in breech face marks occur gradually over time or if more significant changes occur after a random number of discharges. In order to examine this, the running average of the Stepwise Delta 1 comparisons for each firearm was examined separately. The hypothesis was that if the changes in breech face marks occurred through significant changes every few discharges instead of a gradual change, then the match scores would remain relatively constant until a specific Stepwise comparison resulted in a low score representing the significant change. The score would then return to being relatively constant. The majority of the results, however, indicate no pattern at all similar to the results seen in Figure 6. From the lack of pattern, two potential conclusions can be reached. The first is that the changes resulting from a number of discharges do not follow any specific pattern. The second potential conclusion is that there are too many other variables affecting the correlation scores for IBIS® to be able to represent the change in Breech Face Score on the level of individual firearms.

In order to investigate the data in another way, the Stepwise Delta groups for each firearm were pooled together in Figure 7 in order to provide a small sample size for each point. This figure has varying results depending on the variable involved. The Firing Pin Score averages remain steady throughout 200 consecutive test fires with relatively constant variance. The Breech Face Score averages possess a wider range of values with larger variance while still remaining relatively constant over the 200 consecutive test fires. The fact that the scores remain somewhat constant over the 200 consecutive test fires provides evidence that the changes occurring to the breech face over a number of consecutive test fires are gradual. The graphs representing the rank variables show no discernible pattern of value.

Overall the results from Figure 6 and Figure 7 do not provide enough evidence to support the hypothesis that significant changes occur during random discharges while remaining constant otherwise. Due to the constant average score observed in Figure 7a and 7b, it can be seen that the changes in breech face marks observed over a large number of consecutive test fires happens gradually. This would

indicate that these changes are the result of a large number of minute changes that accumulate over time.

4.5 New Firearm's Effect on Breech Face Changes due to Consecutive Test Fires

The fifth objective of this research was to investigate whether the breech face of a new firearm changes more rapidly during the first few discharges of the weapon. This hypothesis is based on the idea that, when a breech face is created, a number of weak striations may be present that will more quickly deform. After a number of discharges, the weak striations would already be deformed and all of the remaining striations would be more robust, resulting in a slower rate of change. If this were true, the expected results would be that the Stepwise comparisons between the first few discharges of the weapons would result in lower scores than those resulting from Stepwise comparison resulting from test observed later in the firearms lifespan.

In order to investigate this, Figure 6 and Figure 7 can again be examined. When examining the Stepwise Delta graphs for each firearm that Figure 6 represents, none of the firearms followed the expected pattern that would indicate that new firearms breech faces change more rapidly. While some firearms did show a rising pattern of increasing correlation scores, it was never significant enough to stand out from the variance that was seen throughout the 200 consecutive test fires observed for each firearm. Furthermore, some firearms show the opposite trend or no trend at all. When observing Figure 7a and 7b, it can be seen that there is no discernible difference between the first few Stepwise comparisons and the rest of the 200 Stepwise comparisons. This indicates that if new firearms are more susceptible to changes in breech face marks, it is not to a significant enough degree that IBIS® can detect it. Furthermore, this might only apply to breech faces that are manufactured in the same manner as the SR9 Ruger. Further research should be conducted to investigate if this is true for firearm breech faces that are manufactured in different manners.

4.6 Worst-Case Scenario Ruger Non-Match vs non-Ruger Known Non-Match

The final objective of this research was to examine if the matching characteristics of different firearms of the same make and model has an effect of Known Non-Match Correlations. This is one of the 'worst-case' scenarios that could occur during casework. In order to examine this, the entire Known Non-Match pool was divided into two groups; one group representing the Known Non-Matches between only Ruger SR9s and the other group representing all of the other Known Non Matches. In Figure 8, we can see that the probability density functions of these two groups can be distinguished

when examining the Breech Face variables. This indicates that IBIS® uses information gathered from the class or sub-class characteristics of the breech face in part of its comparison algorithm. The probability density functions of these two groups showed no separation when examining the Firing Pin variables.

Since class or sub-class characteristics seem to have an effect on the breech face comparison algorithm, Figure 9 was created in order to investigate whether IBIS® could differentiate between Known Matches and Ruger SR9 Known Non-Matches. These results show that in all cases, the Known Match probability density functions can still be distinguished from the Known Non-Matches despite the worst case scenario of matching class characteristics. Furthermore, the Ruger Known Non-Match probability density functions are almost identical to the overall Known Non-Match probability density functions for each variable examined. In Figure 10b below, it can be seen that the Ruger Known Non-Match probability density function is almost identical to the overall Known Non-Match probability density function. This could indicate that the overall Known Non-Match probability density functions are being heavily influenced by the large number of Ruger SR9 Known Non-Matches contained in it. In Figure 10a, we can see that the non-Ruger Known Non-Match probability density function shows much better separation from the Known Match probability density function than the Ruger Known Non-Match probability density function. This indicates that the results of this study represent a worst case scenario and that IBIS® most likely performs better under normal working conditions. In order to examine this further, a study should be performed to evaluate this topic in more detail.

Figure 10: Ruger Known Non-Match vs Other Known Non-Match vs. Overall Density Functions. a) Known Match; b) Known Non-Match

4.7 ROC Curve Comparison: IBIS® vs. 2d Image Comparison

The final objective of this research was to create and compare Receiver Operator Characteristic Curves representing the data obtained from IBIS® and the data obtained from the 2D image comparison. Due to the 2D image comparison algorithms not working correctly, it was not possible to achieve this objective with the results of this research.

5. Conclusions

This study examined a number of variables that could be affecting IBIS® Correlation Scores. The main objective was to examine how the changes observed in the breech face occurring from a number of consecutive discharges effects IBIS® Correlation Score performance. It was found that these changes affected both IBIS® Breech Face and Firing Pin Correlations to different degrees. The Breech Face Correlations are immediately affected over the first few test fires, however, after that the performance decreases at a much slower rate. The initial drop in performance could be caused the small variability between the individual marks between each test fire. This drop in performance levels out because, while present, the variability does not drastically change the overall pattern of individual characteristics being reproduced by the firearm. The Firing Pin Correlations did not exhibit the initial drop in performance and instead showed a very small decrease in performance over time. In general, evidence supports IBIS® ability to accurately differentiate a Known Match from a Known Non-Match to a sufficient degree for it to be used as a screening tool. A secondary objective of this research was to find if the changes in breech face seemed to follow a pattern. Based on the data, changes in breech face marks occur via a number of very minute changes resulting from each test fire that add up over a large number of consecutive test fires. Another secondary objective was to examine if new firearms exhibited a larger rate of change than one that has been 'broken in'. The research indicated that, if new firearms possess a larger rate of change, the difference in rate of change cannot be detected by IBIS®. It is also possible that the difference in the rate of change is small enough to be masked by the variability present between each test fire as mentioned above. The final secondary objective was to examine the 'worst case' scenario involving Known Non-Match comparisons that possess the same class characteristics. The research indicated that the Known Non-Match comparisons with matching class characteristics did exhibit a higher probability density function of score than the other Known Non-Matches; however, it was still distinguishable from the Known Match probability density function. Due to this research, the

following research topics are recommended in the future. Due to the majority of current labs using IBIS® model BRASSTRAX which utilizes 3D topographical information over a 2D image, it is advised that a study similar to this one be repeated on a BRASSTRAX system to ensure that the results are the same across both the 2D and 3D versions of IBIS®. Similarly, this experiment examined a number of SR9 Ruger handguns. Further research should be performed to examine if firearms with breech faces manufactured in a different manor exhibit a stronger or weaker effect on IBIS®'s performance. Research should also be performed that increases the number of consecutive test fires in order to provide a more robust set of data for higher Delta comparisons. Additional research should be conducted to determine the number of test fires required to observe the variability that is present between individual test fires. The final potential area of future research is to design research that will specifically examine Known Non-Matches with matching class characteristics.

In Summation, the following statements can be made. Based on the AUC score of the full Known Match data set, IBIS® is able to designate higher scores to breech face known matches correlations 76% of the time and higher scores to firing pin known match correlations 85% of the time. Based on the AUC scores of the Delta 190+ Known Match data set, IBIS® is still able to designate higher scores to firing pin known match scores 82% of the time however IBIS® is only able to designate higher scores to breech face known match scores 65% of the time. Based on the AUC scores of the individual delta scores observed in Figure 5, the IBIS® correlations scores are affected by the changes in individual marks resulting from consecutive test fires to various degrees. Based on the results observed in Figure 6 and Figure 7, the changes in individual marks resulting from consecutive test fires either occurs in a completely random manner or the pattern of change cannot be detected by IBIS®. Based on the results shown in Figure 6 and Figure 7, the individual marks of new firearms either do not change at a faster rate than when they have been worn in or the change in rate is not detectable by IBIS®. Based on the results observed in Figure 8 and Figure 10, known non-match correlations between cartridge cases from two Ruger SR9 handguns are more likely to receive a higher score then a correlation between a SR9 and a different firearm.

6. References

- [1] Gonzalez, Rafael C.; Woods, Richard E. *Digital Image Processing 2nd e*; Upper Saddle River: Prentice Hall, 2002.
- [2] Scicchitano, Kristine M. *The Effect of Examiner Variation in Cartridge Case Acquisition on IBIS® Correlation Scores and the Ability of the System to Return a True Positive*. M.S. Thesis, West Virginia University, WV, 2011.
- [3] Shem, Robert J.; Striupaitis, Peter P. Comparison of 501 consecutively fired bullets and cartridge cases from a 25 caliber raven pistol. State Crime Lab, Maywood, Illinois.
- [4] Ogihara, Yoshimitsu.; Et al. Comparison of 5000 consecutively fired bullets and cartridge cases from a 45 caliber m1911a1 pistol. Tokyo Met. Crime Lab, Tokyo, Japan. National Research Crime Lab, Tokyo, Japan. Training and Applications Lab, Joliet, Illinois.
- [5] Bernard, Melanie. The Effects of Ammunition Brand and Firearm use on IBIS® Correlation Scores. Can. Soc. Forensic Sci. J. [Online] 2005, 38, 2.
- [6] Saribey, Y. Aylin.; et. al. An Investigation into Whether or Not the Class and Individual Characteristics of Five Turkish Manufactured Pistols Change During Extensive Firing. *Journal of Forensic Science* [Online] 2009, 54, 5.
- [7] Heard, Brian J. Firearms and Ballistics, Examining and Interpreting Forensic Evidence. West Sussex: John Wiley & Sons Ltd, 2008.
- [8] Song J.; et. al., (2012 November). *Establish the “National Ballistics Evidence Search Engine (NBSE)” Based on 3D Topography Measurements on Correlation Cells*. Presented in Forensics@NIST 2012. Symposium conducted at the meeting of the National Institute of Standards and Technology (NIST) Gaithersburg, MD 20899.
- [9] Inc., F.T., *Integrated Ballistics Identification System Training Guide*. 2002. 3.4.
- [10] Yammen, S.; Muneesawang, P. Cartridge case image matching using effective correlation area based method. *Forensic Science International* [Online] 2013, 23.
- [11] Swets, J. A. ROC Analysis Applied to the Evaluation of Medical Imaging Techniques. *Investigative Radiology* [Online] 1979, 14.
- [12] Hanley, J. A.; McNeil B. J. The Meaning and Use of the Area under a Receiver Operating Characteristic (ROC) Curve. *Radiology* [Online] 1982, 143.
- [13] Hanley, J. A.; McNeil B. J. A Method of Comparing the Areas under Receiver Operating Characteristic Curves Derived from the Same Cases. *Radiology* [Online] 1983, 148.