12-14-2015

Spectroscopy and Chemometrics for Cultural Heritage and Forensics: Conservation of Magnetic Tape by Identifying Degraded Polymer Signatures; and Characterizing the Chemiluminescent Reaction of Luminol with Blood

Brianna Marie Cassidy
University of South Carolina - Columbia

Spectroscopy and Chemometrics for Cultural Heritage and Forensics: Conservation of Magnetic Tape by Identifying Degraded Polymer Signatures; and Characterizing the Chemiluminescent Reaction of Luminol with Blood

by

Brianna Marie Cassidy

Bachelor of Science
Murray State University, 2011

Submitted in Partial Fulfillment of the Requirements

For the Degree of Doctor of Philosophy in

Chemistry

College of Arts and Sciences

University of South Carolina

2015

Accepted by:

Stephen L. Morgan, Major Professor

S. Michael Angel, Committee Member

Michael L. Myrick, Committee Member

Claudia Benitez-Nelson, Committee Member

Lacy Ford, Senior Vice Provost and Dean of Graduate Studies

DEDICATION

To: Grandma Delores Applen

ACKNOWLEDGMENTS

"Help others achieve their dreams and you will achieve yours." – Les Brown

Julie Ann Cassidy, it is because of the determination, stubbornness and bravery I inherited from you that this dissertation came to be. You are an incredible mother, role model and friend. I am also grateful for the enduring support and encouragement provided by my family and friends. I warmly thank my fiancé, Zhenyu Lu, who will forever be my best friend and my most admired chemist.

I thank Ms. Lisa Devillez and Dr. Bommanna Loganathan for inspiring me to pursue chemistry and research at a young age. A big thanks goes to my advisor, Dr. Stephen Morgan, my co-advisers Dr. Michael Myrick and Dr. Eric Breitung and my PhD committee for their instruction, guidance and support which led me to reach my goals. I thank Professor Sarah Gutwirth for supporting my passion for art and for teaching me to be comfortable sacrificing something good to create something beautiful.

My colleagues have been a tremendous help in both my research and my moral. A special thanks goes to my lab partners Scott Hoy, Nathan Fuenffinger, Molly Burnip, Kaylee McDonald, and Alyssa Abraham for their sense of humor during difficult times. Stephanie DeJong, Shawna Tazik, Alicia Strange, and Justin Copeland also deserve a large thanks for their intellectual contributions to my work and for their perseverance in getting me out of the lab to have fun every now and then.

ABSTRACT

The first half of this manuscript focuses on the identification of degraded magnetic tape using attenuated total reflectance Fourier transform infrared spectroscopy (ATR-FTIR) and multivariate statistics. For several decades before the digital era, magnetic tape was the dominant audio and visual recording medium. A majority of magnetic tapes contain polyester urethane (PEU) binders, which are known to degrade via hydrolysis, making the retrieval of recorded data difficult and at times impossible. Degraded tapes are currently identified through visual inspection followed by playback on vintage equipment. However, if degraded tapes are played, they are likely to stick and shed onto player guides and heads, resulting in irreversible data-loss. A total of 133 quarter-inch audio tapes were analyzed by ATR-FTIR. Classification of IR spectra in regards to tape playability was accomplished using principal component analysis (PCA) followed by quadratic discriminant analysis (QDA) and K-means cluster analysis. The first principal component suggests intensities at the following wavenumbers to be representative of non-playable tapes: 1730 cm^{-1}, 1700 cm^{-1}, 1255 cm^{-1}, and 1140 cm^{-1}. QDA and cluster analysis both successfully identified 93.78% of non-playable tapes in the calibration set and 92.31% of non-playable tapes in the test set. This application of IR spectra assessed with multivariate statistical analysis offers a path to greatly improve efficiency of audio tape preservation

The second half of this manuscript addresses the wide range of reported detection limits of the forensic luminol test for bloodstains. Luminol (3-aminophthalhydrazide) has been used for blood-stain detection by forensic investigators for over 60 years. When a luminol solution is sprayed onto areas suspected of containing blood, it reacts with the heme moiety of hemoglobin to give a faint bluish-white chemiluminescense. Absolute and relative sensitivities of different luminol formulations have been studied for decades. The range of published luminol detection limits for bloodstains spans nearly five orders of magnitude from 100× to more than 5,000,000× dilute bloodstains. We identify several factors that could affect the response of luminol to dried bloodstains and control them. We obtain a luminol detection limit of ~200,000× diluted blood as an estimate of the best case detection limit. The outcome of this work is a standardized method for measuring the chemiluminescent intensity emitted from the reaction of bloodstains with luminol.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

LIST OF ABBREVIATIONS

ADU ..... Analogue to Digital Units

ATR ..... Attenuated Total Reflectance

FF ..... Fast Forward

FTIR ..... Fourier Transform Infrared Spectroscopy

IR ..... Infrared

LC ..... Library of Congress

LD ..... Limit of Detection

MBRS ..... Motion Picture Broadcasting and Recorded Sound Division

PC ..... Principal Component

PCA ..... Principal Component Analysis

PET ..... Polyethylene Terephthalate

PEU ..... Polyester Urethane

PFA ..... Perfluoroalkoxy

PVC ..... Poly Vinyl Chloride

QDA ..... Quadratic Discriminant Analysis

RW ..... Rewind

SNV ..... Standard Normal Variate

WT% ..... Weight Percent

CHAPTER 1

MAGNETIC TAPES, PLAYABLE OR NOT?

INTRODUCTION

For several decades before the digital era, magnetic tape was the dominant audio and visual recording medium. Recent surveys suggest that over 40 million magnetic tape recordings are held by institutions in the United States alone. The Library of Congress, for example, cares for more than 750,000 tapes. Many tapes are composed of polyester urethane and undergo a degradation process called hydrolysis when exposed to humidity in storage. Hydrolysis causes tapes to become sticky−a phenomenon known all too well by archivists as "sticky shed syndrome" (SSS), which manifests as tapes stick and shed onto playback equipment. The result is permanent tape data loss, impairment of scarce playback equipment, and downtime for equipment cleaning.

Many institutions are in the process of migrating unstable analog tape recordings to digital media. The migration process requires playback of the tape on vintage equipment, digitization of the analog signal, and transfer to a new carrier. This process is significantly hampered by the absence of reliable detection methods for degraded tape, which at present are identified during the actual playback of the tape. Tapes identified as

degraded can often be restored to a playable state by 'baking' at elevated temperatures; however, treatment time for each tape is significant.

Clearly, a non-destructive and rapid tool to detect sticky shed is needed. The Library of Congress (LC) partnered with analytical chemists at the University of South Carolina under the auspices of the Institute of Museum and Library Services. The premise of our research is that construction of such a tool would allow institutions to triage their collection: tapes identified as non-playable can be baked, while tapes identified as playable can be digitized. Recent literature suggested that infrared (IR) light might be used to differentiate degraded from non-degraded magnetic tapes. A beam of IR light is directed on the tape and an absorbance spectrum is recorded which reveals the molecules present. Infrared spectrometry is already widely used in cultural heritage institutions for identification of organic compounds and polymeric materials.

We disseminated a survey to over 50 U.S.-based archives, museums, and libraries to identify the most common tape formats in need of restoration. The responses indicated that quarter-inch reel-to-reel audio tape were in greatest need. Over 100 quarter-inch tapes were acquired from the LC Motion Picture Broadcasting and Recorded Sound Division (MBRS, Culpeper, VA). Each tape was played by an audio engineer using original playback equipment to determine its playability status. IR analysis was then carried out on each tape using a portable instrument (Figure 1.1).

Similarities and dissimilarities among magnetic tape spectra are often found just by looking for one or more absorption peaks associated with specific chemical components. Magnetic tapes, however, are chemically complex, and hydrolysis may not be the only chemical marker of degradation. When a tape begins to degrade, chemical

changes manifest in a "fingerprint" of multiple degradation products that appear in the IR spectrum. A combination of correlated changes is often the key to recognizing significant differences between the spectrum of a tape that is playable, and one that is too sticky to be played. Because IR spectra require assessment at several hundred to several thousand wavelengths, simple visual inspection does not suffice. Our analysis of over 2,000 spectra from more than 100 quarter-inch magnetic tapes has demonstrated that multivariate statistical analysis is able to quickly discriminate the subtle IR patterns of degraded tapes from those of non-degraded tapes. We have also developed a user-friendly software application that enables tape custodians to predict playability status of tapes from IR spectra. Appealing aspects of this approach from the tape conservator's point of view are: (a) the non-invasive nature of the infrared analysis, which does not compromise sound fidelity; (b) IR spectra can be obtained in less than a minute per spectrum; and, (c) a decision to 'bake or 'digitize' is available almost immediately. The combination of FTIR analysis and multivariate statistics has proven to be reliable: a prediction model based on 1900 spectra from 95 representative tapes produced 94% accuracy for prediction of non-playability. When this model was applied to another 760 spectra from 38 entirely different tapes, the prediction accuracy was 92% for non-playability.

The outcome of this IMLS-funded research is that we have developed an effective tool for use by collection custodians to rapidly and non-destructively predict magnetic tape playability. This project will directly impact preservation and digitization work flows at museums, archives, and libraries by providing a validated tool for reliable identification of degraded tapes prior to playing—thus rescuing the nation's recorded heritage from loss.

ACKNOWLEDGEMENTS

This project was supported by Grant LG-06-12-056912 from the Institute of Museum and Library Services to the University of South Carolina. The authors are also grateful for support from the LC by Gene DeAnna and Larry Miller (MBRS) and Fenella France (Preservation Research and Testing), and contributions of Samantha E. Skelton (USC undergraduate researcher, and 2011 Junior Fellows Summer Intern at the Library of Congress), Eric J. Bringley (USC Magellan Scholar and summer 2013 research participant at the Library of Congress), and Linchi Nguyen (summer 2013, ACS Project SEED research participant at the Library of Congress).

Figure 1.1: Infrared analysis of a quarter-inch tape using a portable spectrometer

CHAPTER 2

MINIMALLY INVASIVE IDENTIFICATION OF DEGRADED POLYESTER-URETHANE MAGNETIC TAPE USING ATTENUATED TOTAL REFLECTION FOURIER TRANSFORM INFRARED SPECTROSCOPY AND MULTIVARIATE STATISTICS

ABSTRACT

Audio recordings are a significant component of the world's modern cultural history and are retained for future generations in libraries, archives, and museums. The vast majority of tapes contain polyester urethane (PEU) as the magnetic particle binder, the degradation of which threatens the playability and integrity of these often unique recordings. Magnetic tapes with stored historical data are degrading and need to be identified prior to digitization and/or preservation. We demonstrate the successful differentiation of playable and non-playable quarter-inch audio tapes, allowing the minimally invasive triage of tape collections. Without such a method, recordings are put at risk during playback, which is the current method for identifying degraded tapes. A total of 133 quarter-inch audio tapes were analyzed by attenuated total reflectance Fourier transform infrared spectroscopy (ATR-FTIR). Classification of IR spectra in regards to tape playability was accomplished using principal component analysis (PCA) followed

by quadratic discriminant analysis (QDA) and K-means cluster analysis. The first principal component suggests intensities at the following wavenumbers to be representative of non-playable tapes: 1730 cm^{-1}, 1700 cm^{-1}, 1255 cm^{-1}, and 1140 cm^{-1}. QDA and cluster analysis both successfully identified 93.78% of non-playable tapes in the calibration set and 92.31% of non-playable tapes in the test set. This application of IR spectra assessed with multivariate statistical analysis offers a path to greatly improve efficiency of audio tape preservation. This rapid, minimally invasive technique shows potential to replace the manual playback test, a potentially destructive technique, ultimately allowing the safe preservation of culturally valuable content.

INTRODUCTION

Cultural heritage institutions, such as libraries, museums, and archives with large audio and video collections, hold massive collections of magnetic tape. More than forty-million collection items of recorded sound are held in U.S. archives alone and over 40% are in unknown condition.[1] Many institutions are in the process of digitizing their collections to prevent loss of recorded content due to tape degradation and decreasing availability of functional playback equipment. The digitization process requires playing of tapes using original playback devices, digitization of the analog signal, and transfer to a new carrier.

Magnetic tapes, usually produced with polyester urethane (PEU) binders, were the globally dominant recording medium before the digital era. PEU degrades via hydrolysis, forming products that can make retrieval of recorded data difficult and at times

impossible.[2] Although visual inspection is often attempted to identify degraded tapes, the ultimate test of degradation is playability, usually performed by an audio engineer. Degraded tapes, however, can stick and shed onto player guides and heads, leading to permanent information loss and equipment downtime.[2] A major challenge to the preservation of historically significant sound and video recordings on magnetic tape is the lack of a nondestructive and reliable method for identifying degradation before playing.[2-4] If tape degradation can be reliably identified without passing tapes over multiple guides and play-record heads, degraded tapes can be treated with minimal risk to recordings. The current treatment is a practice commonly referred to as "baking," where tapes are subjected to elevated temperatures (~50 °C) for 8 or more hours.[4,5]

Magnetic tape is a multilayer product consisting of a magnetic layer, a substrate, and an optional antistatic back coating.[6,7] The magnetic layer, which comes into contact with player heads, is about 2−4 μm thick and usually contains Fe2O3 or CrO2 magnetic particles embedded in a PEU binder.3,8 Additives in the magnetic layer may include lubricants to reduce friction during playback, dispersants to improve metal particle distribution, and abrasives to clean machine heads.[2,6,8] The relatively thick substrate (~35 μm), on which the magnetic layer rests, is typically polyethylene terephthalate (PET) and is responsible for the physical integrity of the tape.[2,4,7] A thin back-coating (~1−2 μm) of carbon particles embedded in PEU is often included to dissipate static charge accumulated during playing and rewinding.[7,8]

Availability of components and ongoing process-optimization has resulted in frequent tape formulation changes, sometimes within the same batch and almost certainly within the same model of tape.[7] Information regarding specific formulation components

has generally been unattainable due to lack of transparency or records.[7] Even if formulations were consistently produced within each make and model of tape, manufacturer labels are rarely found on the actual tape. Labels are usually found on the hub or container, though common rehousing practices generally render this information unreliable.

The magnetic data-containing layer is exposed to more mechanical stress during playback than other tape layers.[8] It is also the most complex layer with many possible constituents other than binder and magnetic particles, which make tapes subject to a number of degradation mechanisms.8 Literature agrees that hydrolysis of polyester linkages is the primary degradation mechanism of PEU magnetic tapes (Figure 2.1).[6,9,10] Polyurethane linkages in the PEU binder, and the PET substrate layer, are less sensitive to hydrolysis.[3,9] Degradation contributions from tape components other than the polyester binder, such as magnetic particles[6,11,12] and the back-coat layer,5 have been identified but are not generally regarded as primary degradation paths.[1,9,13,14]

Diverse research studies have addressed detection of tape degradation and its mechanisms. An understanding of degradation processes is required to make predictions of tape life expectancy[9,12,14] and optimal storage conditions.[9,14,10] Reduction of average molecular weight has been observed in artificial aging of tape binder mockups and of PEU elastomers using size exclusion chromatography and other methods. [3,8,11,12,15] Though acidity is a valuable deterioration marker, accurate surface acidity measurement methods for magnetic tape have not surfaced.[8] Tape degradation has been characterized by weight loss after extraction, but results are highly variable.[9,10] Most methods proposed for detection of magnetic tape degradation require destructive sampling, which is

inappropriate for use with unique recordings. The majority of these approaches have been instrumental in improving understanding of tape degradation mechanisms, but tend to be unreliable, require trained technicians, necessitate complex instrumentation, and are time-consuming.

Infrared analysis has been used primarily as a prescreening tool for magnetic-tape samples.[9,10,15] Thiébaut et al. used ATR-FTIR spectroscopy for analysis of tape extracts to identify polymer composition of tapes but suggested the technique unfit to provide further information.[3] However, the ability of ATR-FTIR analysis to provide detailed chemical information on magnetic tape binders without compromising their sound quality has been demonstrated.[2] Furthermore, Hobaica employed ATR-FTIR for analysis of magnetic tape and used spectral information to identify 71% of the degraded quarter-inch tapes in his data set.[4] Although this study showed the potential of ATR-FTIR to differentiate degraded from non-degraded tape, visual inspection of spectra was relied upon to identify and assign absorption peaks for decision making. Because IR spectra have absorbance measurements at thousands of frequencies, visual inspection does not suffice, especially when millions of tapes await inspection. A degraded tape, for example, likely has multiple chemical moieties that correlate with degradation: a "fingerprint" of degradation products manifests itself in the IR spectrum. Thus, a combination of correlated spectral characteristics is key in recognizing significant differences between spectra of playable tapes and spectra of tapes that are too degraded to be played.

Reliable classification of degraded and non-degraded tape by spectroscopy has been elusive due to the variability caused by tape formulation changes and varying levels of degradation present in vintage tapes. The major advantage of using multivariate

statistics to analyze IR spectra is that spectral regions displaying the greatest discrimination for playable and non-playable tapes can be identified and used in the statistical analysis to determine the degradation status of a tape of unknown condition. The objectives of our present research include (1) conducting fundamental studies to validate methods for classification of playable/non-playable magnetic tape using ATR-FTIR spectroscopy and (2) applying multivariate statistical algorithms to facilitate rapid and noninvasive tape evaluation.

EXPERIMENTAL SECTION

Tape Sample Selection. The direct experience with audio tape at the Library of Congress Motion Picture Broadcasting and Recorded Sound Division (MBRS, Culpeper, VA) is that quarter-inch reel-to-reel tape is a significant fraction of collections and is in the most need of preservation due to degradation. For the present study, 133 PEU-based quarter-inch tapes were selected from the MBRS collection which were purchased by the Library of Congress in the 1970s and 1980s. Tapes purchased in this time frame were found by MBRS to be particularly vulnerable to degradation. Though brand and model information were sometimes found on a tape hub or box, this information could not be regarded as a reliable representation of the tape due to common rehousing practices. Only for tapes in original, unopened shrink wrap (which is not the case for any tapes on which information has been recorded) is there a guarantee that company information present on the box and hub is representative of the tape. Of the 133 chosen tapes, 95 were employed

as a calibration set, while the remaining 38 tapes were employed as an independent test set.

Playability Testing. Playback of tapes on vintage tape players is the standard method for differentiating degraded from non-degraded tape. While subjective, manual testing in this fashion is the only current option because other approaches require destructive sampling.[8] All tapes were played by a sound engineer at the MBRS using a Scully 280 tape player (Scully Recording Instruments, Bridgeport, CT). Playability status was determined by attempting to pass the entire tape from one real to the next over six stationary guides (with the read and recording heads removed). Tapes were categorized as non-playable if (a) friction between the tape and player guides slowed or stopped the tape transport; (b) the tape produced squealing noises at any time; (c) the tape exhibited slow speed recovery, indicative of increased friction, between fast forward (FF) and rewind (RW) transitions; or (d) significant tape material sloughed onto player guides. Although tapes exhibit degradation effects during playback in different ways, tapes that showed any of these four behaviors were categorized as non-playable. Tapes were categorized as playable if they exhibited smooth and quiet playback in play, FF, and RW modes without sticking or shedding. Small amounts of edge shed and transfer of magnetic-layer material onto the guides was not considered indicative of a non-playable tape. Videos which exhibit different playability status are included in Appendix A.

Infrared Analysis. ATR FT-IR spectra of the magnetic side of all tape samples were obtained using a Nexus 670 FT-IR running Omnic version 8.2 (Thermo-Nicolet, Madison, WI) equipped with a DTGS detector and a Thunderdome ATR (Thermo

Spectra Tech, Inc., Shelton, CT). The ATR employed a germanium crystal and an incident angle of 45°. The surface of the tape was protected from the metal anvil of the ATR tower by placing Mylar film (~0.1 mm thick) on top of the tape followed by a steel plate (~1 mm thick) to disperse pressure (Figure 2.2). Audio testing confirmed that sound fidelity was not affected by the occasional 2−3 mm diameter smudge observed on tape surface at the ATR pressure point (not shown here).To characterize the degradation state of each of the 133 tapes, 10 spectra were taken at 50 and 100 cm from the beginning of each tape. Replicate spectra were collected after advancing the tape toward the tape hub 3 mm from the previous analysis site. Infrared spectra were acquired over the range of 4000 cm^{-1} to 600 cm^{-1}, with 32 scans at 4 cm^{-1} resolution, without further spectral or ATR corrections. All spectra for the 95-tape calibration model were collected by one analyst in 2011, while the spectra for the 38-tape test set were collected by two other analysts in 2013.

Data Analysis. The data set, consisting of 20 spectra from each of 133 magnetic tapes, was preprocessed and analyzed by multivariate analysis using The Unscrambler X (CAMO Software AS, Oslo, Norway). The 2660 spectra were truncated to range from 1750 to 950 cm^{-1}, corresponding to the spectral region found in previous literature to be descriptive of degradation status for quarter-inch magnetic tape.[4] Spectra were then smoothed using a 13-point Savitzky-Golay fourth order polynomial.[16] The standard normal variate (SNV) transform was applied to remove multiplicative interferences due to scatter and particle size. The SNV transform was applied to each spectrum independently by subtracting its mean absorbance and then dividing each spectral intensity by the standard deviation of the absorbance across the spectrum.[17] Mean

centering was performed after all other preprocessing by subtracting the mean absorbance of the 1900 spectra calibration set from the calibration set itself, as well as from the 760 spectra test set.

Dimensionality reduction of the 95-tape calibration set was achieved by performing PCA on the covariance matrix of the preprocessed 1900 spectra.18−20 The optimum number of principal components needed to describe variation within the 95-tape calibration set was found by performing hold-out-cross-validation. [21] This was done by applying the Kennard−Stone algorithm to select 1140 spectra which were uniformly distributed in the data space of the full 1900 spectra calibration set.[22,23] These 1140 spectra were used as a calibration subset, and the remaining 760 spectra were used as a validation set. The percent variance explained by the calibration subset and the validation set for each of the first 15 PCs was calculated. The calibration subset and validation set explained the most similar amount of variance with five PCs (94.00% and 93.99%, respectively), suggesting that the model is most robust when five PCs are used. Therefore, the first five PCs of the 95-tape calibration set, which account for 94.13% of the total variability within the data set, were employed for classification.

A supervised classification model, where playability information is used to train the model, was built using the first five PCs of the 95-tape calibration set. Quadratic discriminant analysis (QDA), assuming unequal covariance matrices and a multivariate normal distribution for playable and non-playable classes, was used to classify the 95-tape calibration set based on the Mahalanobis distance. The Mahalanobis distance from each sample to both class centroids was calculated, and samples were classified as belonging to whichever group centroid to which they were closest.[18,24] The 38-tape test

set was then projected into the same five dimensional space as the 95-tape calibration set and was also classified using the minimum Mahalanobis distance classifier.

An unsupervised classification model, where playability information is not used to train the model, was also built using the first five PCs of the 95-tape calibration set. K-means cluster analysis assigns an arbitrary centroid location for a user defined number of groups; each spectrum is assigned to the nearest centroid based on squared Euclidian distance.[25] Cluster centroids are recalculated after all spectra have been assigned. Minimization of the sum of squares of distances between spectra and their corresponding centroid and reassignment of classification is iterated until assignments do not change. K-means cluster analysis was carried out, with the number of groups set to two, on the 95-tape calibration set; 50 iterations were sufficient to achieve unchanging cluster assignments. The two cluster centroids determined by K-means cluster analysis of the 95-tape calibration model were used to classify both the 95-tape calibration set and the 38-tape test set by assigning each spectrum to the nearest centroid based on squared Euclidian distance.

RESULTS AND DISCUSSION

Visual Analysis. The potential for differentiating between playable and non-playable tapes using ATR FT-IR is illustrated by Figure 2.3, which shows the average spectra of playable and non-playable tapes from the 95-tape calibration set after a 13-point Savitzky Golay fourth-order polynomial smooth and SNV preprocessing.

Hobaica also analyzed quarter-inch PEU tapes using ATR FT-IR and through visual inspection of the spectra, identified characteristics attributed to non-playable tapes at 1730, 1693, 1364, 1252, and 1138 cm^{-1}.[4] However, there are many differences between playable and non-playable tape spectra, some of which are subtle and not always present in individual spectra.

Principal Component Analysis. PCA of the 95-tape calibration set shows substantial separation between spectra from playable and non-playable tapes along the direction of the first PC, which accounts for 54% of the variability in the data set (Figure 2.4). Clearly, tape playability is a predominant source of variability within tape spectra.

Spectra of playable tapes generally form a tight group within PC space compared to the larger, more spread out group of spectra from non-playable tapes. Variability in the spectra of non-playable tapes may be due to the range of degradation states. Figure 2.4 also shows clusters of spectra from playable tapes in regions where spectra from non-playable tapes tend to reside. This does not necessarily mean these spectra were misclassified as non-playable, since three additional PCs were used for classification.

Even though the same sampling area is never interrogated more than once, spectra taken from the same tape tend to be located within a similar area of PC space. The cluster of playable tape spectra found at coordinates (1, 6) in Figure 2.4 consists of all 20 spectra taken of a single playable tape. The cluster of playable tape spectra found at (4, 4) consists of 40 spectra, representing all replicate spectra for two playable tapes. However, spectra from the same tape do not always reside in one area of PC space. For example, the playable tape spectra found at (0, −2) represent 10 spectra taken 50 cm from the beginning of one tape. The rest of the spectra for the same tape, taken 100 cm from the

beginning of tape, reside near coordinates (−3, 1). Thus, not all tapes are chemically homogeneous throughout their entire length, and spectra from different locations on the same tape may not be classified in the same category.

Principal Component Loading Analysis. Sources of variability contributing to differentiation between playable and non-playable spectra were explored using PCA. The underlying chemical differences responsible for separation of playable tape spectra and non-playable tape spectra can be observed in the eigenvector components for the first PC loading (Figure 2.5). Separation of playable tape spectra from non-playable tape spectra is observed by first looking at Figure 2.4. Non-playable tape spectra tend to reside in the positive direction of the first PC, while playable tape spectra tend to reside in the negative direction of the first PC. For the first PC, eigenvector components which are positive are correlated with non-playable tapes, while eigenvector components which are negative are correlated with playable tapes. As seen in Figure 2.5, the positive components (peaks A, B, and C) are representative of wavenumber regions where hydrolysis products (alcohols and carboxylic acids) absorb, while negative components (D and E) are representative of wavenumber regions where esters absorb. Interestingly, this trend models the process of hydrolysis (Figure 2.1). Hydrolysis of polyester binder has long been agreed as a main degradation process of magnetic tape.[6,9,10]

The carbonyl stretching vibration exhibits the most influential loading (Figure 2.5). However, both esters and their carboxylic acid hydrolysis products contain the carbonyl functional group (Figure 2.1). Therefore, reasons for this feature being correlated with tape degradation were explored using Spartan (Wave function, Inc., Irvine, CA). The IR absorption intensity was calculated for a PEU ester and PEU

carboxylic acid hydrolysis product using density functional theory on a 6-31G basis at the lowest energy molecular geometry. The carbonyl stretch vibrations were found to be inherently less intense for the PEU ester than for the PEU carboxylic acid hydrolysis product. Thus, the ability of the carbonyl stretch absorption to indicate PEU hydrolysis was confirmed.

Supervised Classification. QDA was carried out on the 95-tape calibration set using the first five PCs (Figure 2.6 and Table 2.1). The resulting model is referred to as the playability calibration model for the remainder of the manuscript.

The classification of the playability calibration model is summarized in the form of a confusion matrix (Table 2.1). Results show that 1759 of 1900 spectra were classified in agreement with the status assigned by playability testing, resulting in an overall classification accuracy of 92.58%. Of the 920 spectra representing playable tapes, 840 (91.3%) were correctly classified; of the 980 spectra representing non-playable tapes, 919 (93.78%) were correctly classified.

For five tapes, all 20 spectra were classified differently by QDA than by the playability test. For three tapes, QDA classified spectra from the same tape into seemingly contradictory playable and non-playable categories. Two of these three tapes had all spectra collected 50 cm from the beginning of the tape categorize in one group while the spectra collected at 100 cm categorized in the other. The third tape had only one spectrum categorize differently than the other 19.

QDA was then carried out on the 38 tape test set using the first five PCs of the playability calibration model (Figure 2.7 and Table 2.2). The classification of the

playability test set, based on QDA classification from the group means of the playability calibration model, is summarized in the form of a confusion matrix (Table 2.2). Results show that 650 of 760 spectra were classified in agreement with the status assigned by playability testing, resulting in an overall classification accuracy of 85.53%. Of the 500 spectra representing playable tapes, 410 (82.00%) were correctly classified; of the 260 spectra representing non-playable tapes, 240 (92.31%) were correctly classified.

From 13 non-playable tapes in the test set, one tape was misclassified, having all 20 of its replicate spectra classify as playable. Of the 25 playable tapes in the test set, four tapes were completely misclassified (80 spectra), plus all 10 spectra taken 100 cm from the beginning of another playable tape. Figures 2.6 and 2.7 and Tables 2.1 and 2.2 illustrate the success of training the classification model by optimizing its agreement with playability tests. However, the following expected disagreements between infrared analysis and playability analysis suggest that a classification model may not be optimally trained using the playability test.

As discovered from PCA exploratory analysis (Figure 2.4), all spectra from the same tape do not always reside in a similar region of PC space. Furthermore, spectra from the same tape do not classify into one playability category for four tapes out of the 133 tested. A possible reason for this behavior is that tapes do not degrade homogeneously. Areas which begin to degrade do so at a fast rate since the reaction is autocatalytic,[6,13] yielding blooms of degradation which are initially localized in certain regions of tape. Since infrared sampling took place at two regions along the tape length, spectra taken from the same tape my represent both degraded and non-degraded regions of tape (Figure 2.8).

For 11 tapes out of the 133 tested, all 20 spectra were classified incorrectly. There are several possible reasons for this result: (1) Again, tapes do not degrade homogeneously. Infrared sampling was done within the first 100 cm of the tape, while playability analysis occurred throughout the entire length of each tape (unless tape degradation was severe enough to stop the player). A tape may have a region of degradation encountered only by the playability test, resulting in spectra representative of a non-degraded tape and a playability test representative of degraded tape. (2) Tapes may not show signs of complications during the playability test at an early stage of degradation. However, even at early stages of degradation, multivariate analysis of infrared spectra allows identification of subtle absorbances indicative of degradation. Tapes which do not exhibit playability issues, but are classified by their infrared spectra as non-playable, are likely degrading but are not yet degraded enough to cause playability issues. (3) Complete misclassification of a tape could also be due to the formulation of the tape. Tapes include many varieties of lubricants, dispersants, abrasives, and other components in different relative amounts.[2,6,8] For example, the infrared spectra of a tape may indicate that it is degraded, but the tape may seem playable simply because abundant amounts of lubricant were used in its formulation. (4) Another reason for misclassification of a tape relates to the playability test and the tendency of playability to be effected by the water content of the tape. Adhesion and static friction forces can become substantial between two smooth surfaces, such as between a tape and player guide, in the presence of water.[27] Bushan et al. reported an increase in friction between the tape surface and player guides during playback in environments where temperature and humidity were elevated.[28] Liquid-mediated adhesion, known as stiction, is a function

of the tape water content.[28] Thus, a degraded tape with high water content will exhibit more stiction than a degraded tape with low water content (such as a degraded tape which has recently been baked). This means that a degraded tape categorized as such by its infrared spectra may be found playable if it has low water content.

The above explanations indicate that training a classification model based on infrared spectra using playability as the target could impair model accuracy. For example, if infrared analysis took place on a region of tape which was not yet affected by degradation, but the tape was found to be non-playable within a different region, the model would be trained to categorize playable tape spectra as non-playable tape spectra. Unsupervised classification, however, provides a way to optimize a model without influence of a parallel technique.

Unsupervised Classification. As illustrated by Figure 2.4, PCA analysis of the 95-tape calibration set exhibits separation of the spectral data into two groups over the first PC. This separation is easily observed and is purely due to differences in the chemical composition of the tape samples.

The first five PCs of the 95-tape calibration set were used to carry out K-means cluster analysis. The number of clusters was predefined as two (one cluster for degraded spectra, one cluster for non-degraded spectra). Playability information was not used. The resulting model is referred to as the cluster calibration model for the remainder of this manuscript. The two groups defined by the cluster calibration model will be referred to as degraded tape spectra and non-degraded tape spectra since their designation was based on spectral features alone. The two groups defined by the playability calibration model will

continue to be referred to as playable tape spectra and non-playable tape spectra since their designation was assigned using playability information.

When assigning the first cluster as degraded tape spectra and the second cluster as non-degraded tape spectra, the cluster calibration model agreed with playability designations for 1691 out of 1900 total spectra, giving an overall classification rate of 89.00% (Table 2.3). Furthermore, 100% of individual spectra categorized as non-playable by the playability calibration model classified as degraded by the cluster calibration model.

The cluster calibration model was more conservative in classifying spectra as belonging to non-degraded tape than the playability calibration model. Sixty additional spectra (three total tapes) determined to be playable by the playability and the QDA calibration model were classified as degraded by K-means cluster calibration model. This phenomenon may be due to reasons 2−4 discussed above.

The cluster calibration model was used to classify the 38-tape test set. The playability designation determined by playback agreed with the degradation designation assigned by the cluster calibration model for 620 out of 760 total spectra in the 38-tape test set, giving an overall classification rate of 81.58% (Table 2.4). Again, all spectra assigned as belonging to non-playable tape by the playability calibration model were assigned as belonging to degraded tape by the cluster calibration model. Similar to classification results of the cluster calibration model, more of the 38-tape test set was classified as degraded than was classified as non-playable by the playability calibration model.

Though the unsupervised method of cluster analysis generally agrees less with playback results than the supervised method of QDA, the basis of building a model using purely chemical information without influence from a subjective parallel test may be more desirable. Training an infrared model using playability could render a confused model. Furthermore, the model created using an unsupervised method, without influence of the playability test, has the same ability to recognize non-playable tapes as the supervised QDA method plus the ability to identify playable tapes which show signs of degradation in their spectra and may soon become non-playable.

CONCLUSIONS

Many cultural institutions are in the process of testing, treating, and either digitizing or migrating their audio/visual and data collections held on magnetic tape. A rapid, reliable, minimally destructive, and automated system for identifying degraded tapes is greatly needed. The availability of modeling tools for identifying degraded tapes will increase efficiency in digitization and preservation. While it is not clear how much longer PEU tapes will survive before they are untreatable, the ability to identify the most vulnerable tapes will improve preservation efforts for magnetic tape.

Variance explained by the first principal component is directly related to infrared absorbance changes due to hydrolysis degradation of polyester. QDA prediction of playability demonstrates high potential for ATR-FTIR combined with multivariate statistics to identify degraded tape, identifying 93.78% of tapes deemed non-playable in the calibration model and 92.31% of the tapes deemed non-playable for the test set.

Cluster analysis prediction of playability demonstrates an even higher potential for identification of degraded tape using ATR-FTIR since it uses infrared spectra alone, without influence from the playability test, to determine the chemical degradation status of tapes. Cluster analysis predictions match the ability of supervised QDA predictions in identifying spectra of degraded tape (93.78% for the calibration model and 92.31% for the test set). Furthermore, cluster analysis prediction tends to categorize more playable tape spectra as degraded tape spectra than QDA. This may be because, for example, tapes in the beginning stages of degradation are still playable. Spectra from these tapes are classified as degraded tape spectra by cluster analysis but are trained to be classified as playable tape spectra by QDA. As previously observed (Figure 2.8), tapes do not degrade homogeneously. This means that the ability of ATR-FTIR to identify degraded tape is dependent on the regions of tape analyzed being representative of degraded tape. Thus, more or less infrared analysis may be desirable, depending on the importance of the recording at stake.

In this work, ATR-FTIR with multivariate modeling has proven to be the most accurate, minimally invasive published method in predicting tape playability. The accuracy of this method may be specific within PEU quarter-inch tapes and may not be applicable to other tape formats which are currently under investigation in our laboratories. Tape sample sets from other institutions are also being analyzed to establish the universality of this growing data set.

The method presented in this manuscript has the potential to significantly increase efficiency of digitization workflows and substantially decrease loss of unique recorded sound. Archivists appreciate that the entire preprocessing and statistical procedure can be

automated, allowing them to simply upload spectra of tapes and click a button for a response as to the degradation status of each tape based on a predefined model.

ACKNOWLEDGEMENTS

This project was supported by Grant LG-06-12-056912 from the Institute of Museum and Library Services to the University of South Carolina. The authors thank Bing Gu (Department of Chemistry and Biochemistry, University of South Carolina) for performing Spartan analysis. The authors are also grateful for support from the LC by Gene DeAnna and Larry Miller (MBRS) and Fenella France (Preservation Research and Testing). Brianna Cassidy acknowledges financial support from the Department of Education through GAANN award P200A120075.

REFERENCES

(1) A Public Trust at Risk: The Heritage Health Index Report on the State of America's Collections; Heritage Preservation, Inc.: Washington, DC, **2005**.

(2) Gómez-Sánchez, E.; Simon, S.; Koch, L.-C.; Wiedmann, A.; Weber, T.; Mengel, M. ATR-FTIR Spectroscopy for the Characterization of Magnetic Tape Materials. *E-Preservation Sci*. **2011**, *8*, 1−9.

(3) Thiébaut, B.; Vilmont, L.; Lavédrine, B. Characterization of U-Matic Videotape Deterioration by Size Exclusion Chromatography and Pyrolysis Gas Chromatography/Mass Spectrometry and the Role of Adipic Acid. *J. Cultural Heritage* **2009**, *10*, 183−197.

(4) Hobaica, S. Analysis of Audio Magnetic Tapes with Sticky Shed Syndrome by ATR-FTIR. *J. Appl. Polym. Sci*. **2013**, *128*, 1962−1973.

(5) Richardson C. A. Process for Restoring Magnetic Recording Tape Damaged by Sticky Shed Syndrome. U.S. Patent 6,797,072, September 28, 2004.

(6) Edge, M.; Allen, N. S.; Hayes, M.; Jewitt, T. S. Degradation of Magnetic Tape: Support and Binder Stability. *Polym. Degrad. Stab*. **1993**, *39*, 207−214.

(7) Hess, R. L. Tape Degradation Factors and Challenges in Predicting Tape Life *Assoc. Record. Sound Collect. J*. **2008**, *34*, 240−274.

(8) Thiébaut, B.; Vilmont, L.; Lavedrine, B. Report on Video and Audio Tape Deterioration Mechanisms and Considerations about Implementation of a Collection Condition Assessment Method; Report Deliverable D6.1; PrestoSpace: Paris, France, 2006.

(9) Cuddihy, E. F. Aging of Magnetic Recording Tape. *IEEE Trans. Magn*. **1980**, *16*, 558 −568.

(10) Bertram, H. N.; Cuddihy, E. F. Kinetics of the Humid Aging of Magnetic Recording Tape. *IEEE Trans. Magn.* **1982**, *18*, 993−999.

(11) Bradshaw, R.; Bhushan, B.; Kalthoff, C.; Warne, M. Chemical and Mechanical Performance of Flexible Magnetic Tape Containing Chromium dioxide. *IBM J. Res. Dev*. **1986**, *30*, 203−216 V.

(12) Yamamoto, K. Watanabe, H. A Kinetic Study of Hydrolysis of Polyester Binder in Magnetic Tape. In Proceedings of the Fourth NASA Goddard Conference on Mass Storage Systems and Technologies, College Park, MD, March 29−30, 1995.

(13) Brown, D. W.; Lowry, R. E.; Smith, L. E. Prediction of the Long Term Stability of Polyester-Based Recording Media; Report Number NBSIR 84-2988; National Bureau of Standards: Washington, DC, 1984.

(14) Nakamae, K.; Yamaguchi, K.; Asaoka, S.; Karube, Y.; Sudaryanto. Lifetime Expectancy of Polyurethane Binder as Magnetic Recording Media *Int. J. Adhes. Adhes*. **1996**, *16*, 277−283.

(15) Brown, D. W.; Lowry, R. E.; Smith, L. E. Kinetics of Hydrolytic Aging of Polyester Urethane Elastomers. *Macromolecules* **1980**, *13*, 248−252.

(16) Savitzky, A.; Golay, M. J. E. Smoothing and Differentiation of Data by Simplified Least Squares Procedures. *Anal. Chem*. **1964**, *36*, 1627−39.

(17) Barnes, R. J.; Dhanoa, M. S.; Lister, S. Standard Normal Variate Transformation and De-trending of Near-Infrared Diffuse Reflectance Spectra *Applied Spectroscopy*. **1989**, *43*, 772−777.

(18) Krzanowski, W. J. Principles of Multivariate Analysis: A User's Perspective, rev. ed.; Oxford University Press: New York, 2000.

(19) Rencher, A. C. Methods of Multivariate Analysis, 2nd ed.; John Wiley & Sons: New York, 2002.

(20) Morgan, S. L.; Bartick, E. G. Discrimination of Forensic Analytical Chemical Data Using Multivariate Statistics. In Forensic Analysis on the Cutting Edge: New Methods for Trace Evidence Analysis; Blackledge, R. D., Ed.; John Wiley & Sons: New York, 2007; pp 331−372.

(21) Esbensen, K. H.; Geladi, P**.** Principles of Proper Validation: Use and Abuse of Re-Sampling for Validation. *Journal of Chemometrics* **2010**, *24*, 168−187.

(22) Kennard, R. W.; Stone, L. A. Computer Aided Design of Experiments. *Technometrics* **1969**, *11*, 137−148.

(23) Daszykowski, M.; Walczak, B.; Massart, D. L. Representative Subset Selection. *Anal. Chim. Acta* **2002**, *468*, 91−103.

(24) Mahalanobis, P. C. On the Generalized Distance in Statistics. *Proc. Natl. Inst. Sci. India* **1936**, *2*, 49−55.

(25) MacQueen, J. Some Methods for Classification and Analysis of Multivariate Observations. Proceedings of the Fifth Berkeley Symposium on Mathematical Statistics and Probability, Berkeley, CA, December 27, 1965.

(26) Smith, B. C. Infrared Spectral Interpretation: A Systematic Approach; CRC Press: Boca Raton, FL, 1999.

(27) Bhushan, B. Adhesion and Stiction: Mechanisms, Measurement Techniques, and Methods for Reduction. *J. Vac. Sci. Technol., B.* **2003**, *21*, 2262−2296.

(28) Ambekar, P. P.; Bhushan, B. J. Effect of Operating Environment on Head–Tape Interface in a Linear Tape Drive. *Journal of Magnetism and Magnetic Materials* **2003**, *261*, 277−294.

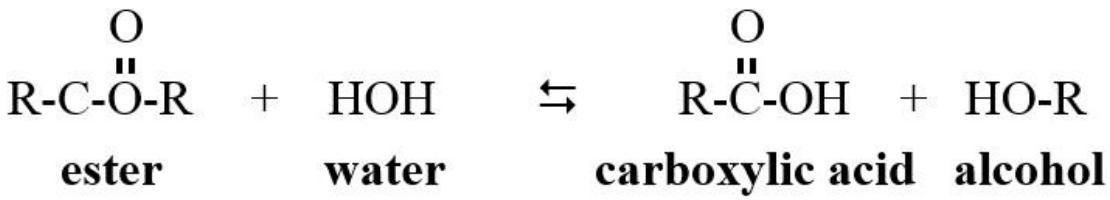

Figure 2.1. Hydrolysis of polyester is an autocatalytic reaction, in which production of carboxylic acid causes further hydrolysis.[3,6,13] Hydrolytic degradation of polyester binder results in formation of low molecular weight products which are sticky in nature and cause tribological issues in magnetic tape.[2,9]

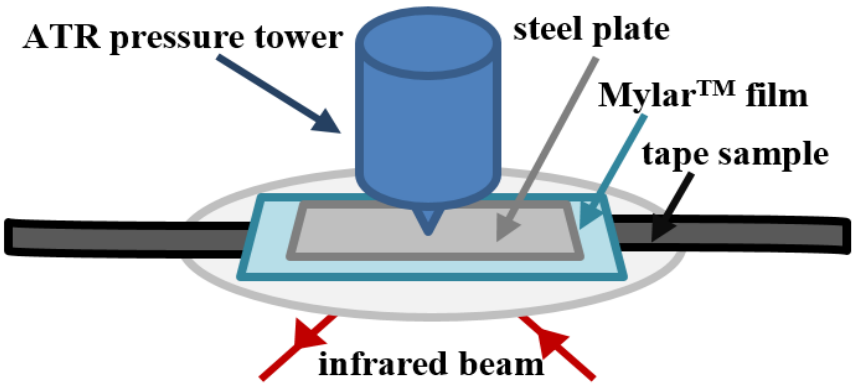

Figure 2.2. Minimally invasive ATR sampling of magnetic tape.

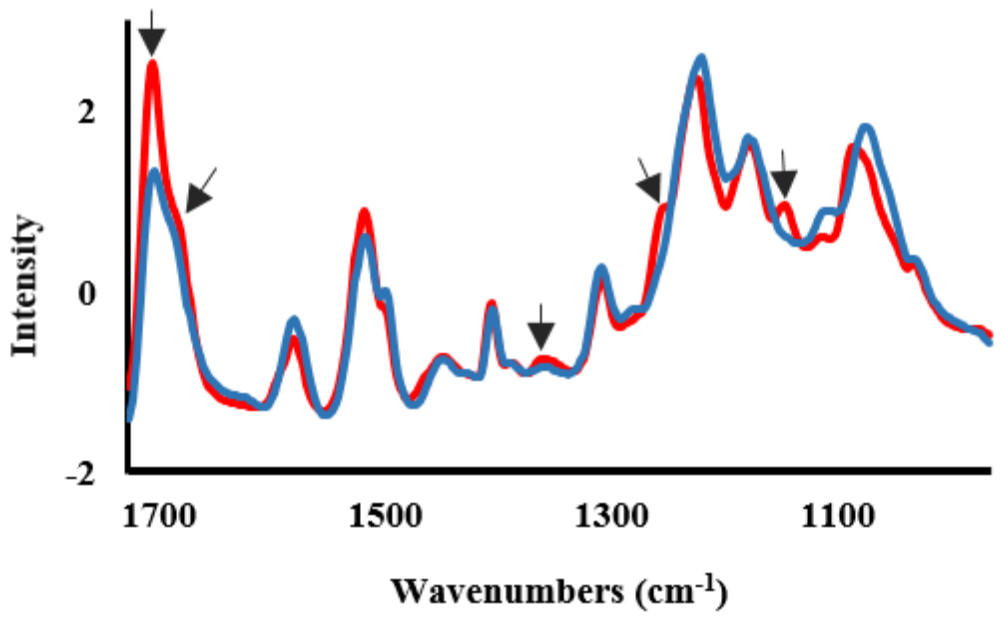

Figure 2.3. Average preprocessed spectra of 51 playable tapes (blue) and 44 non-playable tapes (red). Arrows indicate aforementioned wavenumbers found by Hobaica to be representative of non-playable tape.

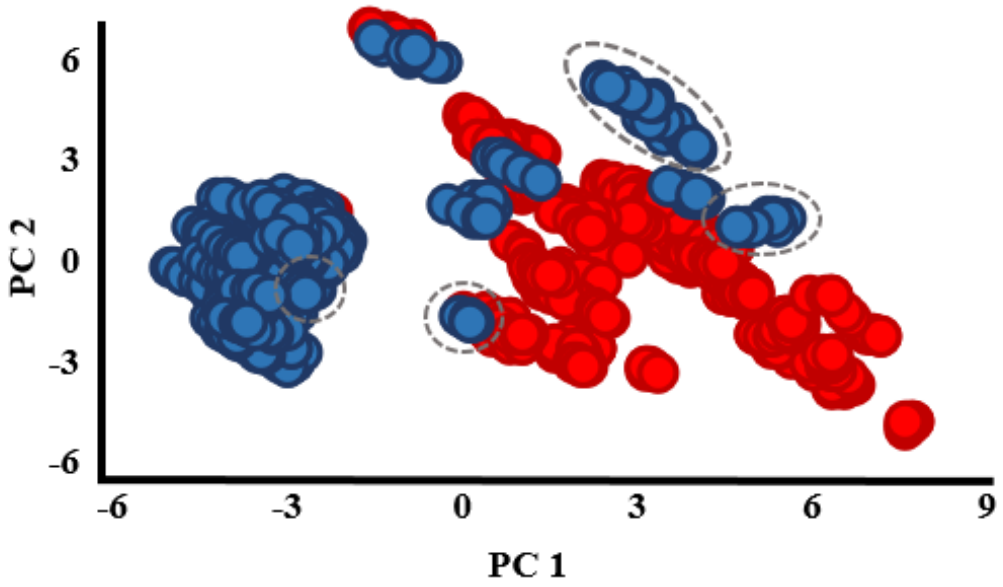

Figure 2.4. Projection of the 95-tape calibration set (1900 spectra) into the space of the first two principal components. Each of the 920 blue data points represent a spectrum from a playable tape and each of the 980 red data points represent a spectrum from a non-playable tape. PC 1 and PC 2 account for 54% and 20%, respectively, of the total variability within the data set. Data points outlined in gray ellipses are referred to later.

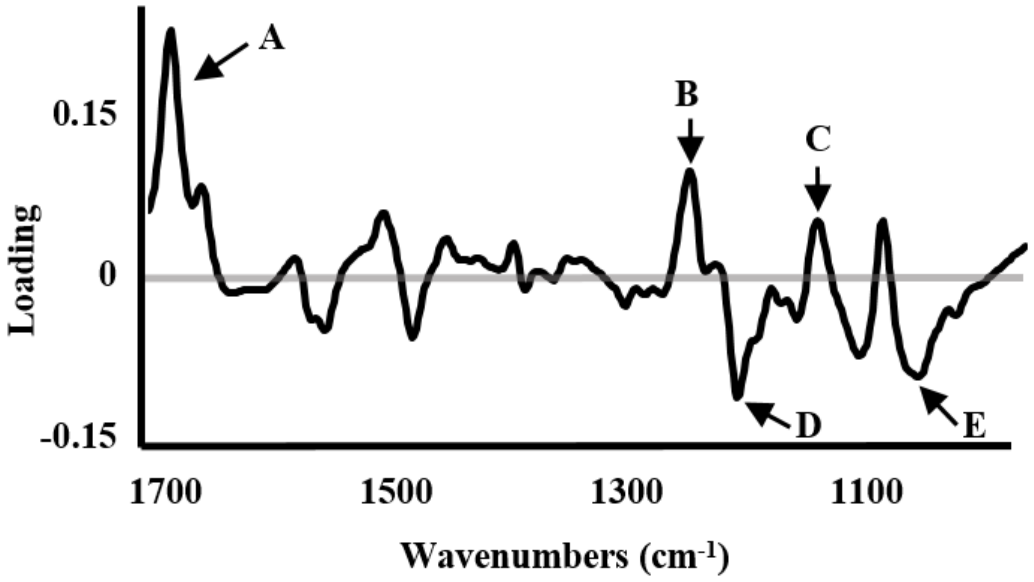

Figure 2.5. Principal component 1 shows chemical differences between playable and non-playable tapes of the 95-tape calibration set. Spectral regions representative of non-playable tape are related to positive loadings: (A) due to C=O stretching vibrations at 1700 cm^{-1} to 1730 cm^{-1}; (B) due to the C−O stretching of carboxylic acid at 1255 cm^{-1}, and (C) due to the C−O stretching of alcohol vibrations at 1140 cm−1. Spectral regions representative of playable tapes exhibit negative loadings: (D) due to C−C−O ester stretching at 1210 cm^{-1}; and (E) due to ester O−C−C stretching at 1050 cm^{-1}. Assignments were based on ref 26. Infrared assignments in regard to magnetic tape in addition to those presented here can be found in previous studies.[2,4]

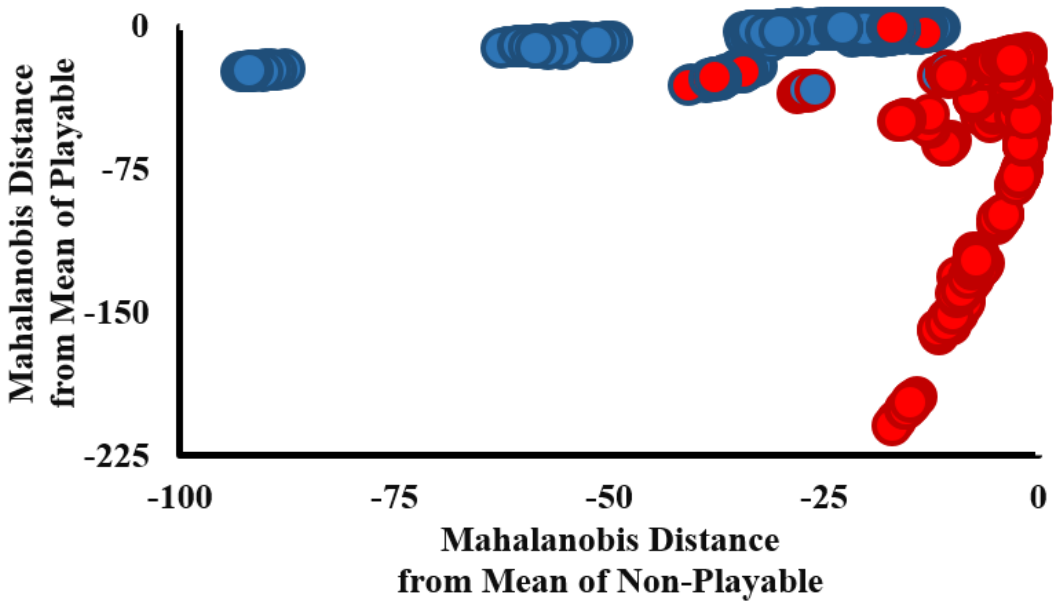

Figure 2.6. QDA results of the playability calibration model. All 1,900 spectra are represented by their Mahalanobis distance from the group means of playable and non-playable spectra. Each of the 920 data points with blue centers represent a spectrum from a playable tape and each of the 980 data points with red centers represent a spectrum from a non-playable tape. The 999 data points with red borders were classified by QDA as non-playable tape spectra, while the 901 data points with blue borders were classified by QDA as playable tape spectra.

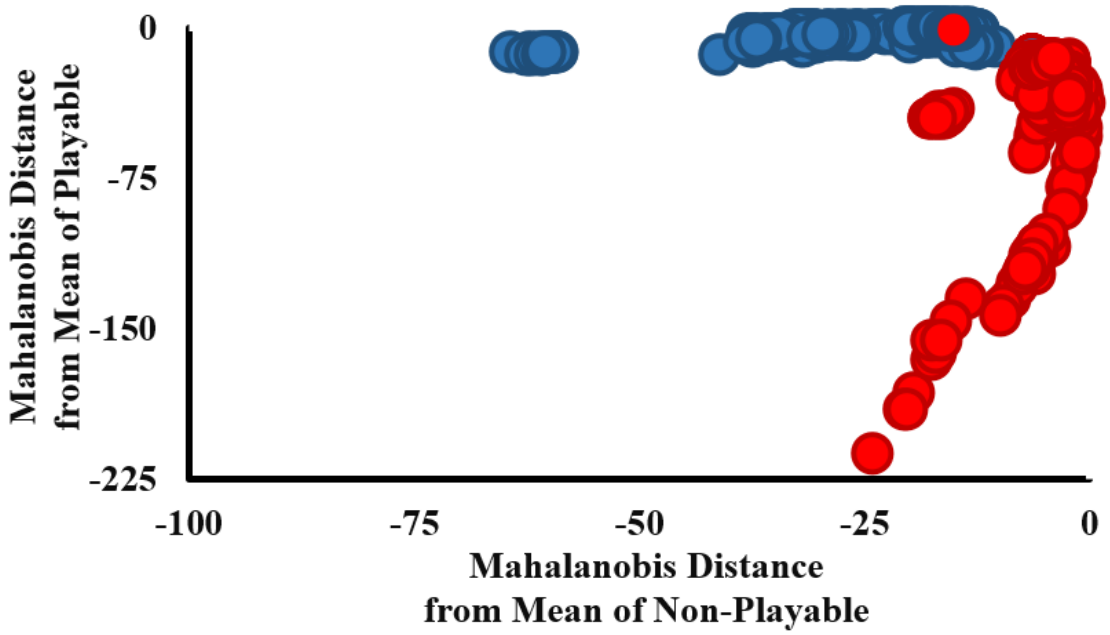

Figure 2.7. QDA results of the playability test set classified using the playability calibration model. All 760 spectra are represented by their Mahalanobis distance from the group means of playable and non-playable spectra of the playability calibration model. Each of the 500 data points with blue centers represent a spectrum from a playable tape; each of the 260 data points with red centers represent a spectrum from a non-playable tape. The 330 data points with red borders were classified by QDA as non-playable tape spectra, while the 430 data points with blue borders were classified by QDA as playable tape spectra.

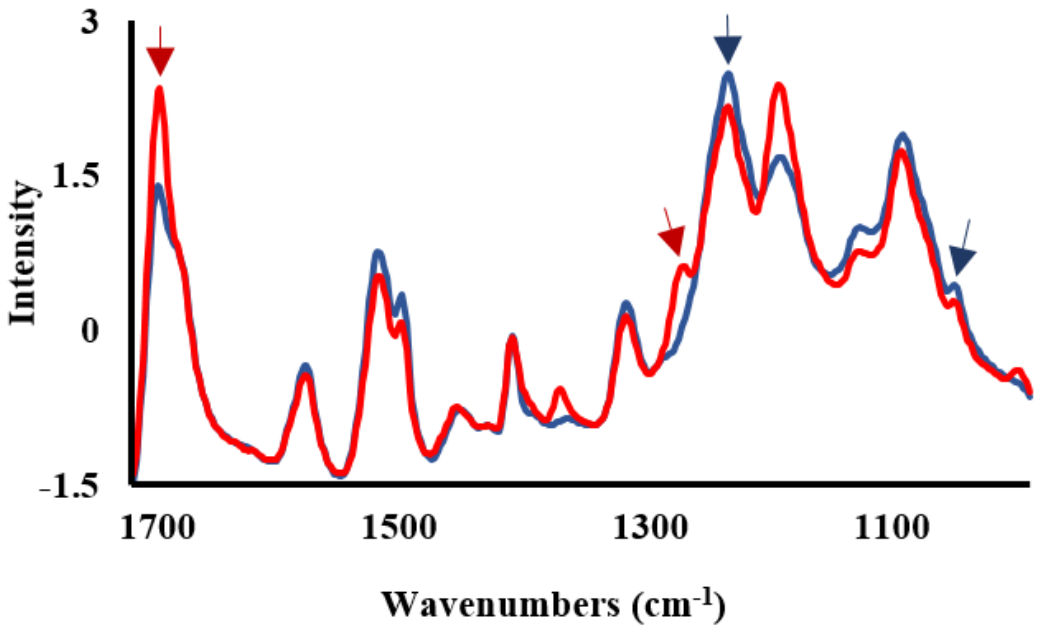

Figure 2.8. Two preprocessed spectra from the same tape. The red spectrum, taken about 50 cm from the tape beginning, was categorized as belonging to a non-playable tape and has strong features at 1700 and 1255 cm^{-1} (indicated with red arrows) found earlier to be representative of non-playable tapes (Figure 2.5). The blue spectrum, taken about 100 cm from the tape beginning, was categorized as belonging to a playable tape and has more prominent features at 1050 and 1210 cm^{-1} (indicated with blue arrows) than the red spectrum. These wavenumber regions were found earlier to be representative of playable tape (Figure 2.5). This figure illustrates that tapes undergo inhomogeneous degradation.

Table 2.1. QDA results vs playability results of the 95-tape calibration set.

Playability Calibration Model (1900 Spectra)	QDA Classification	
Playback Classification	Playable	Non-Playable
Playable	840/920= 91.30%	80/920= 8.70%
Non-Playable	61/980= 6.22%	919/980= 93.78%

Table 2.2. QDA results vs playability results of the 38-tape test set.

Playability Test Set (760 Spectra)	QDA Classification	
Playback Classification	Playable	Non-Playable
Playable	410/500= 82.00%	90/500= 18.00%
Non-Playable	20/260= 7.69%	240/260= 92.31%

Table 2.3. Cluster analysis results *vs*. playability results of the 95-tape calibration model.

Cluster Calibration Model (1900 Spectra)	Cluster Classification	
Playback Classification	Cluster 1	Cluster 2
Playable	772/920= 83.91%	148/920= 16.09%
Non-Playable	61/980= 6.22%	919/980= 93.78%

Table 2.4. Cluster analysis results of the cluster test set *vs*. playability results of the 38-tape test set.

Cluster Test Set (760 Spectra)	Cluster Classification	
Playback Classification	Cluster 1	Cluster 2
Playable	380/500= 76.00%	120/500= 24.00%
Non-Playable	20/260= 7.69%	240/260= 92.31%

.

CHAPTER 3

OPTIMUM CASE DETECTION LIMIT OF THE FORENSIC LUMINOL TEST FOR LATENT BLOODSTAINS

ABSTRACT

The luminol test has been used for over 60 years by forensic investigators for presumptive identification of blood and visualization of blood splatter patterns. Multiple studies have estimated the limit of detection (LD) for bloodstains when luminol is employed, with results ranging from 100× to 5,000,000× dilute. However, these studies typically have not identified and controlled important experimental variables which may affect the luminol LD for bloodstains. In this manuscript, we report an optimum case luminol LD for bloodstains at ~200,000× diluted blood after controlling numerous experimental factors associated with the blood-luminol reaction that could degrade detection.

INTRODUCTION

The initial task of a forensic crime scene investigator is to recognize items that might have evidentiary value. Blood is among the most commonly encountered bodily fluids encountered by forensic investigators. The advent of trace DNA amplification has increased the importance of blood detection. However, if bloodstains have been diluted by deliberate washing of the substrate or by environmental exposure to rain or submersion in water, detection can be compromised.[1] Latent stains, those invisible to the naked eye, may result if only trace amounts of blood remain. If bloodstains are present but not identified, crucial evidence is overlooked as illustrated by the Damilola Taylor case.[1] Therefore, if there is reason to believe blood might be present, a presumptive test is often performed, with positive results followed by confirmatory tests.[2-4]

Luminol (3-aminophthalhydrazide) is among the most sensitive blood detection reagents available for forensic investigation and has been employed for decades.[2,3,5] Luminol solutions for bloodstain detection are typically alkaline and contain hydrogen peroxide as an oxidizing reagent. Ferric heme groups in blood catalyze oxidation of luminol by decomposition of hydrogen peroxide present in the luminol solution.[5] The chemiluminescent reaction path of luminol has been studied for over 50 years. Studies have described the general reaction mechanism to involve the oxidation and excitation of luminol resulting in the excited state dianion intermediate, 3-aminophthalate, that upon return to ground state emits a broad spectrum of light centered around 425 nm (Figure 3.1).[5,6-8] White, et al. identified 3-aminophthalate as the light emitting species in the luminol reaction by matching its fluorescence spectrum to the chemiluminescent spectrum

of luminol.[7] However, further details involving the light-emitting pathway of luminol remain speculative and intermediates have not been completely characterized.[5,9-11]

Luminol applied to bloodstains has been shown not to negatively affect subsequent DNA analysis of bloodstains.[6, 12-15] However, luminol has been reported to denature blood enzymes, with consequent effects on the biochemical profile.[11] Grispino and Laux, *et al*. emphasize that patent bloodstains should never be contaminated with presumptive blood detection reagents.[11, 15]

Absolute and relative sensitivities of presumptive tests for blood have been studied for over 60 years.[16] Some studies compare non-luminol based techniques to luminol-based techniques,[2,13,14,16,17] while other studies compare the sensitivity of different luminol formulations.[9,18] However, results are inconsistent. For example, Bluestar® (a commercialized luminol formula) was reported to outperform a luminol solution prepared according to a police department's crime lab protocol.[18] Patel, *et al*. tested five luminol formulations and found Bluestar® Magnum to have greater sensitivity than other formulations on both porous and non-porous surfaces.[9] Seashols, *et al*. modified a luminol solution previously suggested by Grodsky[16] and found it to perform similarly to Bluestar® for all tested cases, except when tested on linoleum.[17] The range of published luminol LDs for bloodstains spans nearly five orders of magnitude from 100× to more than 5,000,000× dilute bloodstains.

Lack of agreement among estimated LDs of presumptive tests for bloodstains is of concern to the forensic community.[9,13, 19,20] Cox, *et al.* attribute the large range of reported LDs to variations in substrates, sample preparation methods, reagent concentrations, and result interpretations.[20] Recently, DeJong, *et al*. suggested

inconsistencies of reported LDs are due to the absence of blank measurements, lack of quantitative detection methods, and lack of data validation, among other reasons.[19] In this manuscript, we identify several factors that could affect the response of luminol to dried bloodstains and control them – the methods by which blood is measured, blood dilutions are made, replicate bloodstain samples are made, luminol is prepared, luminol is stored, luminol is applied to bloodstains, and the method by which chemiluminescent response is detected. We estimate a best case LD of ~200,000× diluted blood on cotton. The outcome of this work is a standardized method for measuring the reaction of bloodstains with luminol to enable accurate and reproducible determination of LDs. The techniques provided here can also be used in further experiments to show how variations in the factors we have controlled degrade the luminol LD.

EXPERIMENTAL SECTION

Substrate Preparation. The substrate used in this study, 8 oz. 100% Cotton Twill Wingfoot (Milliken, Oakbrook, IL), was cut into 5″ × 5″ swatches. Each swatch was sonicated in 100 mL of methanol (ACS grade, Sigma Aldrich, CAS: 67-56-1, St. Louis, MO) for one hour and hung to dry overnight in a fume hood. These samples are simply referred to as "cotton substrate" in the following text.

Stain Barrier Application. Determination of a LD depends heavily on the ability to create reproducible samples. Previous luminol LD studies report depositing bloodstain solutions in measured aliquots. However, the more dilute a blood solution, the further the solution spreads when applied to a substrate (Figure 3.2, top). Van Dalan, *et al*. deposited

300 µL of varying dilutions of porcine blood solutions on cotton and reported substantial increases in spot size with dilution increase.[21] This phenomenon, which complicates calculation and relationship of the mass of blood solids per amount of substrate and introduces an element of randomness from sample to sample, has been previously confirmed using mouse blood dilutions on the same cotton substrate used in this study.[19] Implementation of a stain barrier, made of polyvinyl chloride (PVC) cement (Oatey®, regular clear PVC cement #31012, Cleveland, OH) was used to create a defined area in which liquids may spread.[22] The cotton substrate was first placed in a 3″ embroidery hoop. Caps taken from screw thread vials (Fisherbrand®, CAT NO 03-339-21F, Pittsburgh, PA) were placed on each side of the fabric so that the open ends of the caps face each other with the fabric sample in-between. A C-clamp (Grizzly Industrial®, Model: G8094, Springfield, MO) was used to securely hold the caps in place. Using a cotton swab, one coat of PVC cement was applied liberally around the caps on each side of the cotton substrate and allowed to dry for at least one hour before the C-clamp and caps were removed. Further description of the method for applying stain barriers can be found in the Supporting Information. Stain barriers limit the effect that blood dilution has on spreadability, and enhance sample reproducibility (Figure 3.2, bottom).

Bloodstain Preparation. Blood tends to stick on the inside of pipette tips and container surfaces. When pipetting whole blood, or concentrated blood solutions, a film of blood or blood solution is left in the pipette tip. This coating effect results in less than anticipated amounts of blood or blood solution being deposited, which in turn results in creation of blood solutions which are less concentrated than intended. Medina, *et al*. report that inaccuracies due to the coating effect may be as high as 12% of the intended

volume when working with whole blood.[23] Simple options exist to combat such coating effects when working with blood: hydrophobic containers should be used when creating and storing blood solutions (such as silanized glassware or Teflon® fluoroplastic resin containers); and, and the reverse pipette method should be used whenever pipetting volumes of blood or blood solutions.[24,23]

Mouse blood was received in a BD Vacutainer® vial (K2 EDTA, REF 367862, Franklin Lakes, NJ) from the Department of Animal Resources (University of South Carolina, Columbia, SC). Seashols, *et al*. reported that EDTA-treated blood is equivalent to untreated blood for forensic luminol experiments.[17] Blood was received within one hour of g extraction from live mice to avoid clotting. Blood dilutions were made by mass on a Mettler-Toledo AG204 analytical balance, based on the absolute gravity of water at 25.5°C (0.9969 g/mL) and blood (1.0595 g/mL).[25]

Blood solutions were prepared in 20 mL perfluoroalkoxy (PFA) bottles (Jenson Inert Products, item # GBT5009-20, Coral Springs, FL) using HPLC grade water (Fisher Scientific, CAS: 7732-18-5, Fair Lawn, NJ) to avoid possible contaminants (e.g., iron). A primary 94.168× diluted blood solution (nominally 100×) was made and sonicated for one hour, after which blood solutions of 22,585×, 26,101×, 29,360×, 37,288×, 44,634×, 58,524×, 97,852 × and 186,035× dilute were produced within two hours of the blood being extracted from live mice (Figure 3.3). The peculiarity of these dilution factors stems from the fact they were created by mass on an analytical balance. More regular dilution factors could be produced if more time was allotted for depositing masses of liquids closer to the back-calculated masses from blood solutions of 25,000×, 28,571×, 33,333×, 40,000×, 50,000×, 66,667×, 100,000× and 200,000× dilute. Approximate

masses were deposited instead to allow fast conversion of whole blood into solution, which ultimately results in a reduction of complications due to blood clotting. To reduce the presence of blood clots, and increase homogeneity of the solutions, all blood solutions were sonicated for one hour directly after preparation.

Blood solutions were deposited onto fabric substrates immediately after sonication. Each fabric swatch was suspended in an embroidery hoop placed on top of an open container, ensuring that the substrate is level and that deposited solutions do not interact with surfaces other than the substrate. If the sample were allowed to rest on a surface, not only could it be easily contaminated, but liquid deposited on the sample may be pulled to unquantifiable areas due to capillary action between the sample and surface. Four replicate bloodstains for each of the eight concentrations were created by depositing 100 µL of blood solution into the center of the stain barriers using the reverse pipette method.[24] Blank samples were made by depositing 100 µL of HPLC grade water into eight different stain barriers. All samples were allowed to dry overnight under ambient conditions. No visible indication of blood could be seen on the cotton substrate after the diluted blood solutions were applied.

Because bloodstains were allowed to dry and were analyzed as dried solids, bloodstain amounts were converted from dilution factor (a unit of measurement appropriate for liquid solutions) to weight percent (WT%). Blood is composed of approximately 20% solid material.[26] Thus, mass in grams of liquid whole blood in 100 µL of each solution was multiplied by 0.2 to calculate the mass of blood solids on deposited on cotton substrate. The mass of cotton substrate on which the blood solutions were applied was found by cutting and weighing the fabric from stain barriers. The

average mass of the cotton substrate inside the stain barriers was 0.0979 g (standard deviation = 0.0028 g, n = 5). The resulting weight percent blood solids on each dried cotton substrate was 9.58, 8.29, 7.37, 5.80, 4.85, 3.70, 2.21, and 1.16×10^{-4} respectively, for 22,585×, 26,101×, 29,360×, 37,288×, 44,634×, 58,524×, 97,852× and 186,035× diluted blood.

Luminol Preparation. Bluestar® was purchased from Arrowhead Forensics (Lenea, KS). Two tablets from each vial were placed in 250 mL of HPLC grade water (Fisher Scientific, CAS: 7732-18-5, Fair Lawn, NJ) in an Erlenmeyer flask equipped with a stir bar set on low. The stir bar was stopped once all four tablets dissolved (approximately 6 min.). Luminol solutions have a finite time in which they are prescribed to be used.[5] Most literature suggests that luminol solutions not be used after three hours from the time of their preparation. We found that the luminol solution degrades within the first three hours after solution preparation (Figure 3.4). In this study, to insure a consistent 'effective age' of the luminol of 15 min. throughout the series of experiments, we stored a single luminol solution in 1 mL aliquots in cryogenic vials (Globe Scientific, CryoClear Model no. 3002-500), and submersed the vials simultaneously in liquid nitrogen 10 min. after the time tablets had dissolved. Because the frozen solutions were applied 5 min. after melting, luminol solutions were effectively 15 min. old at time of deposition for all analysis.

Luminol Application. Another source of luminol LD variability stems from the luminol application step. No study to our knowledge has reported quantitative results regarding the effect of adding more or less detection agent to bloodstains, which is likely to have an effect on the response. Any instance where a spray bottle is used, even if the

amount dispensed is measured,[27] does not guarantee that all liquid dispensed interacts with the bloodstain. Some studies controlled the amount of detection agent applied to a bloodstain by applying the detection agent dropwise directly onto the stain.[3,11] However, without control of the area to which a liquid spreads on a substrate, it is not likely that the entire applied luminol solution interacted with the entire bloodstain (Figure 3.5, top). The stain barrier technique introduced in this manuscript ensures the full amount of pipetted luminol solution is constrained to interact with the entire bloodstain (Figure 3.5, bottom).

Blood-Luminol Detection. Most studies employed visual observation as the method of blood-luminol response detection. [2,3,9,13,14,16,17,27,28] However, visual observation of luminol response is highly subjective. Finnis, *et al.* and Lytle, *et al*. did use a camera to record blood-luminol chemiluminescent responses, but did not take measurements from their photographs, and responses were ultimately assessed qualitatively by visual observation.[1,29] Implementation of a quantitative method for detecting blood/detection-agent response is required to attain reliable, reproducible sensitivity results, as well as to provide fair comparison of the performance of different presumptive bloodstain detection agents. Our procedure detects the luminol chemiluminescent response using long exposure raw images obtained using a CCD camera from which quantitative pixel intensity information is extracted (Figure 3.6).

A Nikon D80 CCD camera (Nikon Americas Inc., Melville, NY) was equipped with a Nikkor 50 mm f/1.8D AF lens (Nikon), kept at maximum zoom for the entirety of the experiment. The camera was focused on samples using the adjustable height optical support rod secured to an optical breadboard (Newport, Irvine, CA) for the entirety of the experiment (Figure 3.6). A remote control (NEEWER, RS-60E3, China) was used to set

the exposure time to 5 min. and to permit camera initiation without direct camera contact. All photographs were saved in 12-bit raw format. The darkroom was completely sealed from light and all indicator lights on the camera, camera remote, *etc*., were covered with blackout tape to prevent stray light affecting results. The camera ISO (similar to gain) was set to 500. The exposure time (or integration time) was optimized by comparing the most concentrated dilution chemiluminescent response to the blank chemiluminescent response using 30 sec. exposures over a 10 min. time period. The time at which the most concentrated dilution and the blank chemiluminescent response intensities converged, 5 min., was selected as the exposure time to capture periods of the reaction where bloodstains differ in chemiluminescent response from the blank. Since each 30 sec. exposure raw image requires 70 additional sec. to save, one 5 min. exposure was employed to observe the response in each experiment reported.

Measurement of the Blood-Luminol Response. For each analysis, bloodstain samples were suspended in an embroidery hoop placed on top of an open container to ensure the cotton substrate was level for luminol deposition and that the sample did not come into contact with any surface. The order in which samples were analyzed was randomized with the exception of the blanks, which were deliberately analyzed at even intervals spaced through the entirety of the experiment. Additionally, replicate samples were never analyzed sequentially.

A reference photo was taken in lit conditions with an ISO setting of 100 and an exposure time of 1/10 sec. at the start of each bloodstain analysis to enable stain location in subsequent photos. Three background photos were then taken in complete darkness with the same camera settings used for sample analysis (ISO: 500, exposure: 5 min.). A

cryogenic vial of luminol was then removed from the liquid nitrogen and thawed in a 1,000 mL room temperature water bath. A stir bar was used to move the vial, cap-side-down, into a vortex which allowed steady observation of the melting process (Figure 3.7). The melting process can only be observed if the cryogenic vials are transparent and the luminol solution level is high enough for the last bit of frozen luminol to be observed outside of the opaque cap of the vial. Immediately after the entire luminol solution thawed, it was removed from the water bath and a timer was set for 5 min. The vial was wiped dry, the cap was slightly loosened to relieve any pressure, and was placed inside a dark container. Four min. after the luminol thawed, 100 µL of luminol solution was extracted reverse-pipette style from the cryogenic vial using a micropipette. One hand held the micropipette in position over the stain, while the other hand held the camera remote. The only light present at this time came from a lamp placed next to the sample which was plugged into a power strip on the floor. The lamp was turned off by toggling the power strip on the floor to off by foot just before 5 min. had passed since the luminol solution thawed. At the 5 min. mark, a 5 min. exposure 500 ISO photograph was initiated, and the luminol was deposited onto the sample in a steady controlled fashion using the reverse pipette technique in complete darkness.[25]

Data Analysis. All photos were converted from the Nikon raw format (.NEF) to Matlab® compatible .DNG format using Adobe® Digital Negative Converter (version 1.3, custom compatibility: uncompressed). After photos were imported into Matlab® (Mathworks, Natick, MA), the reference photo was used to locate the bloodstain and to create a mask which ensured that only the bloodstain portion of the image was measured. Since luminol emits light in a band around 425 nm,[5] only blue pixels were considered for

intensity information. The three background photo pixel intensities previously taken were averaged and the resulting background pixel intensities were subtracted from the sample photo pixel intensities. The sum of resulting pixel intensities from the bloodstain location on the sample photo was then calculated as a measure of the quantitative chemiluminescent response. A more detailed description of our photo-analysis method is provided within the Matlab® code provided in the Supplemental Information.

RESULTS AND DISCUSSION

A total of 40 samples were analyzed, four replicates for each of eight bloodstain dilutions plus eight blank replicates. Bloodstain dilutions of 22,585×, 26,101×, 29,360×, 37,288×, 44,634×, 58,524×, 97,852× and 186,035× were converted to measurements of weight percent bloods solids, resulting in weight percentages of 9.58, 8.29, 7.37, 5.80, 4.85, 3.70, 2.21, and 1.16×10^{-4} respectively.

The eight blanks had an average intensity of 2.01×10^{5} ADU and a standard deviation of 5.76×10^{4} ADU. The limit of detection was calculated by setting the predicted response from a quadratic model fitted to the calibration data equal to the average of the blanks plus three times the standard deviation of the blanks, and solving for the weight percent blood solids LD. The limit of detection (using Bluestar®) on the cotton substrate was calculated to be 9.16×10^{-5} weight percent blood solids. This value translates to a deposition of 100 µL of a 235,974× diluted blood solution onto an area of cotton substrate weighing 0.0979 g. Given the relative uncertainty from the experimental

variables under control, and uncertainties in the statistical prediction of the LD, this value might be reasonably rounded to a ballpark estimate of 200,000× diluted blood.

The quadratic model in Figure 3.8 exhibited no significant lack fit, whereas a first-order model did exhibit significant lack of fit.[30] The curvature in the data also suggests that the luminol reaction is of higher order with respect to weight percent blood solids. However, the luminol reaction rate order cannot be resolved because the role of the catalyst and the reaction intermediates have not been completely characterized.[5] Therefore, since determination of a rate order is beyond the scope of this study, a relationship between weight percent blood solids and chemiluminescent response intensity was defined based solely on experiment results.

The heteroscedasticity observed in Figure 3.8 may be due to the effects of residual blood clotting. Sonicating blood solutions tended to reduce the presence and severity of blood clots, but did not alleviate the issue. Increased variation is observed in replicate sample intensities of higher blood content, possibly because interaction and clotting of blood before deposition is more likely in concentrated blood solutions. The more blood clots present in solution, the less homogeneous the solution, and the greater the variability of replicate bloodstains made using that solution. However, because the LD was calculated from the variability of the blanks, and the estimation of the limit of detection is in the region of low level replicates which exhibit near constant variability, statistical adjustment for heteroscedasticity was not required.

The most dilute bloodstains in our study (1.16×10^{-4} WT% blood solids) fall just inside our calculated luminol DL of 9.16×10^{-5} WT% blood solids. One of the four replicates of the lowest weight percent blood solids had a chemiluminescent intensity

which exceeded the estimated Bluestar® LD chemiluminescent intensity (3.73×10^5 ADU). Since samples having the lowest weight percent blood solids fall just inside the calculated luminol LD, over half might be expected to have a response greater than 3.73×10^5 ADU (false positive outcomes). However, the full distribution of intensities is not defined with as few as four replicates; additional low-level replicate measurements might clarify this point.

Additional to weight percent blood solids, the luminol LD is presented as dilution factor to allow comparison to previously reported luminol LDs. Before comparing luminol LDs in terms of dilution factor, a discussion of possible misconceptions related to reporting luminol LDs in terms of dilution factor is required. The dilution factor unit of measurement does not accurately represent the sample being analyzed for several reasons: When LDs are presented as dilution factors the amount of blood deposited goes unaccounted for. For example, a bloodstain created by depositing a 100 µL aliquot of a 100× diluted blood solution cannot readily be differentiated from a bloodstain created by depositing a 1 mL aliquot of 100× diluted blood solution, especially if the amount of blood solution applied is not provided. Both stains were created using the same blood dilution factor, but differ by three orders of magnitude in the solutions spread to a larger area of substrate than more concentrated blood solutions of the same volume (Figure 3.2, top). Therefore, if a 100 µL aliquot of 100× diluted blood solution spreads to a smaller area of substrate than a 100 µL aliquot of 1,000× diluted blood solution, the concentration relationship between these two bloodstains can no longer be accurately described. Likewise, if a 100 µL aliquot of 100× diluted blood solution spreads to a smaller area when deposited on cotton substrate than when deposited on silk substrate,

the bloodstains cannot necessarily be described as having equal concentrations. These factors should be kept in mind when comparing reported luminol LDs in terms of dilution factor.

Converting reported LDs in terms of dilution factor to weight percent blood solids would provide a more accurate comparison, but this it is not possible because previous studies did not report the amount of substrate to which each blood solution spread. Some studies, especially those which soak substrates in blood solutions, do not report the resulting amount of blood solution present on the substrate. In attempt to account for missing information regarding the amount of substrate to which blood solutions spread, DeJong, *et al*. estimated the area by depositing 100 µL aliquots of different blood solution concentrations onto cotton fabric and measured the area to which each solution spread. Although DeJong, *et al*. confirmed that stain area increases with dilution factor, they were unable to visualize the area to which blood solutions of 100× dilute or higher spread, even on white cotton. The luminol LD in terms of dilution factor reported here (235,974×), is more than twice the dilution factor for which the fabric area was able to be estimated in the DeJong, *et al.* study.[19] Additionally, we have found that, if the same volume of a blood solution is applied to two different white cotton substrates (*e.g*., cotton t-shirt *vs*. cotton slacks, the stain spreads to different-sized areas (results not shown here). This effect further complicates the accurate estimation of fabric involved with bloodstains of previous studies. For these reasons, reported LDs in terms of dilution factor were not converted to weight percent blood solids. Deviations in previously reported LDs compared to LD reported here are accounted for, in part, by deviations introduced by reporting LDs in terms of dilution factor.

Reported luminol LDs that are of lower concentration and of higher concentration than our reported 235,974× diluted blood might be accounted for by the potential effects of several experimental factors. (1) More or less than 100 µL of blood solution (the amount used here) was used in other studies. Some studies which reported more sensitive luminol LDs submersed the substrate in blood solutions, and thus likely resulted in more than 100 µL of analyte (blood solution) being involved in the analysis.[16,29] Other studies, which reported less sensitive luminol LDs, only applied 50 µL of blood solution for analysis.[1,17] (2) A more, or less, sensitive luminol formulation was employed. Differences in reported luminol LDs have previously been said to be the result of differences in reagents and reagent concentrations.[9,13,20] It is interesting to note that all reports of more sensitive luminol LDs, that we have found, employed luminol solutions other than the one used in this work (Bluestar®).[2,3,16,29] (3) The luminol solution was applied to bloodstains at a 'younger' or 'older' age (after solution preparation) than 15 min. Grodsky, *et al*., who reported a luminol LD of 5,000,000× diluted blood, emphasized the importance of applying luminol solution as soon as possible after creation.[16] If the age of luminol solution is as crucial as described by Grodsky, *et al*., then the 15 min. old luminol solution used on our study may have resulted in a degraded luminol LD. It is unlikely, however, that a 15 min. difference in solution age would result in over an order of magnitude decrease in LD. Some studies which report less sensitive luminol LDs do not mention the age of the luminol solution at time of deposition,[13,14] others admit to using luminol solutions for several hours after being prepared.[1] The relationship between luminol solution age and blank chemiluminescent response is shown in Figure 3.4. However, further studies which include bloodstain samples additional to blanks is may be

useful in further demonstrating the effect of luminol age. (4) Misinterpretation of luminol responses may result in overestimated luminol LDs if blanks have not been incorporated into the experiment, or if blank responses are similar to responses of dilute bloodstains. As seen in Figure 3.8, even blank samples result in a low intensity light emission, which are observable by the naked eye. In fact, the luminol solution itself was observed emitting low intensity light during our study. Therefore, if all bloodstain samples incorporated in an experiment were reported to have produced an observable chemiluminescent response, and no indication is given of blank sample analysis,[16,29] it is likely that the intensity said to have been observed at extreme dilutions was that initiated by the substrate, or that of the background reaction of the luminol solution itself. Furthermore, even when blanks are measured, it is difficult to visually differentiate between blank samples and heavily-diluted samples: a quantitative method of detection is more appropriate for determining LDs.

Additional reasons exist which may explain less sensitive luminol LDs than the LD reported in this study. Deviations from the experimental procedure which would result in a less sensitive luminol LD are extensive, but the most evident include the following effects. (1) Blood tends to stick to glassware and pipette tips. No study, to our knowledge, reports the use of silanized glassware or fluoroplastic resin containers for blood solution storage. The effect of blood adsorbing to container sides would result in a more dilute than expected solution concentration, which ultimately results in the production of bloodstains that are less concentrated than intended. (2) If the forward pipette method was used instead of the reverse pipette method[24] when measuring aliquots of whole blood, up to 12% of the intended volume may not have been delivered.[23] This

coating effect results in the creation of less concentrated blood solutions than intended and less than intended amounts of blood being deposited onto substrate. (3) Bloodstains created on fabrics which are not suspended from surfaces likely contain less blood solution than intended, due to the absorption of blood to the surface underneath the substrate.[1] If the surface under the substrate is non-porous, capillary action may pull the blood solution to areas less accessible to applied luminol, reducing the chemiluminescent response.

The objective of the research presented in this manuscript is to introduce a new experimental procedures capable of eliminating as many potential sources of experimental variability as possible, permitting accurate interrogation of the influence of selected factors. This study focuses on the effect that bloodstain concentration has on the chemiluminescent response of luminol, ultimately resulting in an optimum case luminol LD. Our results suggest that investigators may not be able to detect bloodstains of concentrations lower than 9.16×10^{-5} WT% blood solids or 235,974× diluted blood using the Bluestar® luminol technique. Furthermore, this study demonstrates that the intensity of the luminol chemiluminescent response is dependent on the amount of blood present, suggesting that even the most faint luminol responses merit further attention from investigators. Experimental techniques introduced here allow independent exploration of variables which may affect the chemiluminescent response and, ultimately, the practical utility of luminol at crime scenes. Similar further experiments might be performed to characterize quantitatively the effects of luminol solution formulation, luminol solution concentration, luminol solution age, substrate effects, bloodstain age, and bloodstain *vs.* interferant response, individually or in unison.

CONCLUSIONS

Detection of bloodstains at crime scenes has been widely studied for decades. Unfortunately, the plethora of previous literature on the subject of sensitivity and comparison of presumptive tests for bloodstains fails to agree and often does not provide conclusive results.[9,13,19,20] Contradictions in reported LDs are rooted in variation introduced by inaccurate blood volume measurements, irreproducible bloodstain preparation methods, neglect of luminol age effects, uncontrolled luminol application, substrate effects, subjective detection techniques, and generally insufficient experimental descriptions.

The experimental procedure described herein maximizes variable control, enabling calculation of an optimum-case Bluestar® LD of 9.16×10^{-5} WT% blood solids. This result may be compared to previous literature once translated to a blood dilution factor of 235,974×. Our results demonstrate that luminol response is dependent on the amount of blood present and that even the weakest response to luminol application merits consideration.

The outcome of this work is a standardized method for measuring the chemiluminescent intensity emitted from the reaction of bloodstains with luminol to maintain accurate and reproducible determination of LDs. Furthermore, our approach may be useful for accurate comparisons between different luminol formulations and between different types of detection agents. Some may question whether the experimental control demonstrated in this work is relevant to the use of bloodstain detection agents in the field, where a high degree of variable control is not permitted. Without control of experimental parameters in the laboratory, variables which affect the potential of presumptive bloodstain

test methods remain largely unknown, and comparisons required to establish new, more powerful detection methods are simply impossible. Without such controlled experimentation, forensic investigators are not knowledgeable of the limitations, or the true utility, of techniques being used to detect precious bloodstain evidence.

ACKNOWLEDGEMENTS

This project was supported by Award No. 2011-IJ-CX-K055 from 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 are those of the author(s) and do not necessarily reflect those of the Department of Justice.

REFERENCES

(1) Finnis, J.; Lewis, J.; Davidson, A. Comparison of Methods for Visualizing Blood on Dark Surfaces. *Sci. Justice.* 2013, *53*, 178-186. DOI:10.1016/j.scijus.2012.09.001

(2) Castello, A.; Alvarez, M.; Verdu, F. Accuracy, Reliability, and Safety of Luminol in Bloodstain Investigation. *Can. Soc. Forensic Sci. J.* 2002, *35*, 113-121.

(3) Webb, J. L.; Creamer, J. I.; Quickenden, T. I. A Comparison of the Presumptive Luminol Test for Blood with Four Non-Chemiluminescent Forensic Techniques. *Luminescence* 2006, *21*, 214-220. DOI:10.1002/bio.908

(4) Creamer, J. I.; Quickenden, T. I.; Apanah, M. V.; Kerr, K. A.; Robertson, P. A Comprehensive Experimental Study of Industrial, Domestic and Environmental Interferences with the Forensic Luminol Test for Blood. *Luminescence* 2003, *18*, 193-198. DOI: 10.1002/bio.723

(5) Barni, F.; Lewis, S. W.; Berti, A.; Miskelly, G. M.; Lago, G. Forensic Application of the Luminol Reaction as a Presumptive Test for Latent Blood Detection. *Talanta* 2007, *72*(*3*), 896–913. DOI:10.1016/j.talanta.2006.12.045

(6) White, E. H.; Zafiriou, O.; Kagi, H. H.; Hill, J. H. M. Chemiluminescence of Luminol: The Chemical Reaction. *J. Am. Chem. Soc.* 1964, *86*, 940–941.

(7) White, E. H.; Zafiriou, O.; Kagi, H. H.; Hill, J. H. M. Chemiluminescence of Luminol and Related Hydrazides: The Light Emission Step. *J. Am. Chem. Soc.* 1964, 86, 941–942.

(8) Soderquist, T. J.; Chesniak, O. M.; Witt, M. R.; Paramo, A.; Keeling, V. A.; Keleher, J. J. Evaluation of the Catalytic Decomposition of H_2O_2 Through Use of Organo-Metallic Complexes - A Potential Link to the Luminol Presumptive Blood Test. *Forensic Sci. Int.* 2012, *219*, 101-105. DOI:10.1016/j.forsciint.2011.12.005

(9) Patel, G.; Hopwood, A. An Evaluation of Luminol Formulations and Their Effect on DNA Profiling. *Int. J. Legal Med.* 2013, *127*, 723-729. DOI 10.1007/s00414-012-0800-9

(10) Barnett, N. W.; Francis, P. S. Encyclopedia of Analytical Science, 2005, 506-545.

(11) Grispino, R. R. The Effect of Luinol on the Serological Analysis of Dried Human Bloodstains. *J. Crime Laboratory Digest* 1990, *17*, 13-22.

(12) Hochmeister, M. N.; Budowle, B.; Baechtel, F. S. Effects of Presumptive Test Reagents on the Ability to Obtain Restriction Fragment Length Polymorphism (RFLP) Patterns from Human Blood and Semen Stains. *J. Forensic Sci.* 1991, *36*(3), 656-661.

(13) Tobe, S. S.; Watson, N.; Daeid, N. N. *J.* Evaluation of Six Presumptive Tests for Blood, Their Specificity, Sensitivity, and Effect on High Molecular-Weight DNA. *J. Forensic Sci.* 2007, *52*, 102-109. DOI: 10.1111/j.1556-4029.2006.00324.x

(14) Budowle, B.; Leggitt, J. L.; Defenbaugh, D. A.; Keys, K. M.; Malkiewicz, S. F. The Presumptive Reagent Fluorescein for Detection of Dilute Bloodstains and Subsequent STR Typing of Recovered DNA. *J. Forensic Sci.* 2000, *45*, 1090-1092.

(15) Laux D. L. Effects of Luminol on the Subsequent Analysis of Bloodstains. *J. Forensic Sci.* 1991 *36*, 1512–1520. DOI: 10.1520/JFS13171J

(16) Grodsky, M.; Wright, K.; Kirk, P. L. Simplified Preliminary Blood Testing—An Improved Technique and a Comparative Study of Methods. *J. Crim. L. Criminology & Police Sci.* 1951, *42*, 95-104.

(17) Seashols, S. J.; Cross, H. D.; Shrader, D. L.; Rief, A. A Comparison of Chemical Enhancements for the Detection of Latent Blood. *J. Forensic Sci.* 2013, *58*, 130-133. DOI:10.1111/j.1556-4029.2012.02259.x

(18) Dilbeck, L. Use of Bluestar Forensic in Lieu of Luminol at Crime Scenes. *Journal of Forensic Identification* 2006, *5*, 706-756.

(19) DeJong, S. A.; Lu, Z.; Cassidy, B.; O'Brien, W.; Morgan, S.; Myrick, M. Detection Limits for Blood on Four Fabric Types Using Infrared Diffuse Reflection Spectroscopy in Mid- and Near-Infrared Spectral Windows. *Anal. Chem.*, 2015, *87*(17). DOI: 10.1021/acs.analchem.5b01825

(20) Cox, M. A Study of the Sensitivity and Specificity of Four Presumptive Tests for Blood. *J. Forensic Sci.* 1991, *36*, 1503-1511.

(21) van Dalen, G. Protein on Cloths: Evaluation of Analytical Techniques. *Appl. Spectrosc.* 2000, *54*, 1350-1356. DOI: 0003-7028/00/5409-1350$2.00/0

(22) Cassidy, B. M.; Lu, Z.; Witherspoon, K. A.; Bensussan, A.; Martin J.; DeJong, S. A.; O'Brien, W.; Morgan, S. L.; Myrick, M. L. A Reproducible Sample Preparation Method for Quantitative Stain Detection. Provisional Patent USC-467-P (1146), April 25, 2015.

(23) Medina, F.; Cheong, V.; Peck, C.; Bensinger, T. A. Imporved Method for Using Eppendorf Pipettes for Accurate Delivery of Blood. *Clin. Chem.* 1977, *23*, 1188-1189.

(24) Thermo Scientific Pipetting Guide, Tips for Good Laboratory Pipetting. 1514170-04 Thermo Fisher Scientific/TopnovaOy/PIPGUIDE-LCP-0407-04. http://fpb.case.edu/smartcenter/docs/SpitCamp/Pipetting%20Guide_Thermo%20Scientific_25440.pdf

(25) van Slyke, D. D.; Phillips, R. A.; Dole, V. P.; Hamilton, P. B.; Archibald, R. M.; Plazin, J. Calculation of Hemoglobin from Blood Specific Gravities. *J. Biol. Chem.* 1950, *183*, 349-360.

(26) Dust, J. M.; Grieshop, C. M.; Parsons, C. M.; Karr-Lilienthal, L. K.; Schasteen, C. S.; Quigley, J. D., III; Merchen, N. R.; Fahey, G. C, Jr. Chemical composition, protein quality, palatability, and digestibility of alternative protein sources for dogs. *J. Anim. Sci.* 2005, *83*, 2414-2422.

(27) Middlestead, C.; Thornton, J. Sensitivity of the Luminol Test with Blue Denim. *J. Forensic Sci.* 2010, *55*, 1340-1342. DOI:10.1111/j.1556-4029.2010.01427.x

(28) Garofano, L.; Pizzamiglio, M.; Marino, A.; Brighenti, A.; Romani, F. A Comparative Study of the Sensitivity and Specifity of Luminal and Fluorescein on Diluted and Aged Bloodstains and Subsequent STRs Typing. *Int. Congr. Ser.* 2006, *1288*, 657-659. DOI:10.1016/j.ics.2005.10.048

(29) Lytle, L. T.; Hedgecock, D. G. Chemiluminescence in the Visualization of Forensic Bloodstains. *J. Forensic Sci.* 1978, *23*, 550-562.

(30) Deming S. N.; Morgan, S. L. The Use of Linear Models and Matrix Least Squares in Clinical Chemistry, *Clin. Chem.* 1979, *25*(6), 840-855.

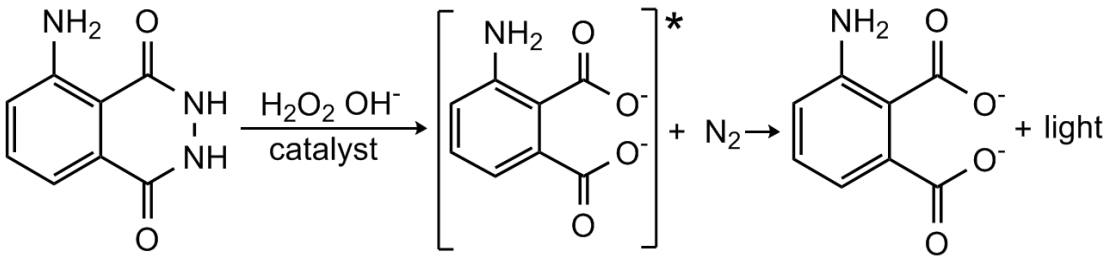

Figure 3.1. The luminol light-emitting reaction involves oxidation of luminol and excitation of an intermediate species which emits light upon returning to ground state.

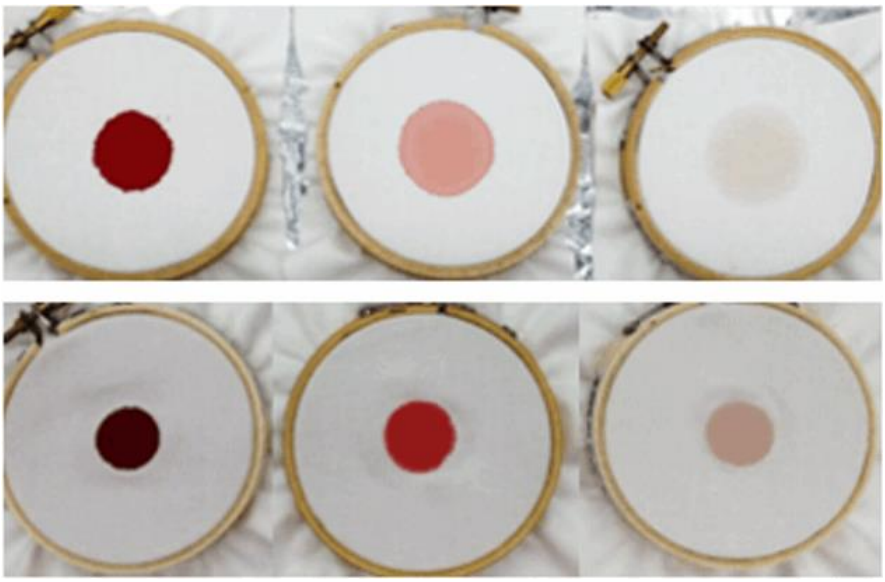

Figure 3.2. Blood solutions (left to right: whole blood, 10× diluted blood and 100× diluted blood) deposited as separate 100 µL aliquots on cotton substrate suspended in 3″ embroidery hoops. Top: The diameter of the resulting stain is dependent on the dilution applied, which ultimately compromises the ability to compare quantitatively one stain dilution to another. Bottom: Blood solutions are restricted by PVC stain barriers to occupy the same, reproducible area of fabric.

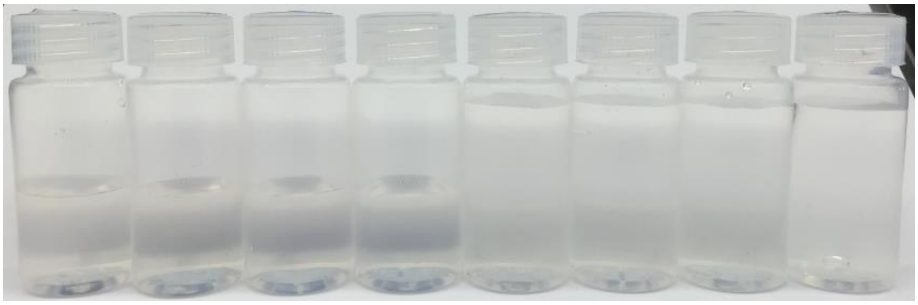

Figure 3.3. Blood solutions (left to right: 22,585×, 26,101×, 29,360×, 37,288×, 44,634×, 58,524×, 97,852× and 186,035× diluted blood) in PFA containers.

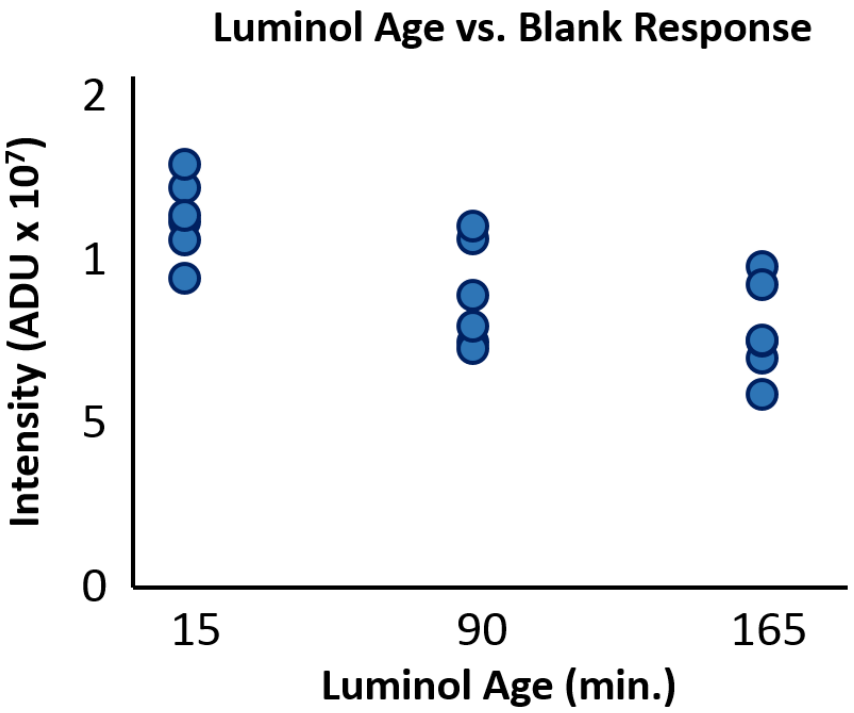

Figure 3.4. Effect of luminol age on blank intensity. White cotton Fruit of the Loom® T-shirts (Bowling Green, KY, RN 13765, 0FLA04) that had been sonicated in methanol for one hour, had stain barriers applied, and had100 µL of HPLC grade water applied to each were used as blanks. Five blanks were analyzed for each luminol solution age. When comparing 15 min. to 165 min. old luminol solution responses, the degradation in luminol response is more than 30%. The average response intensity for 15 min. old luminol on white cotton Fruit of the Loom® T-shirt fabric was 1.13×10^7 with a standard deviation of 1.20×10^6.

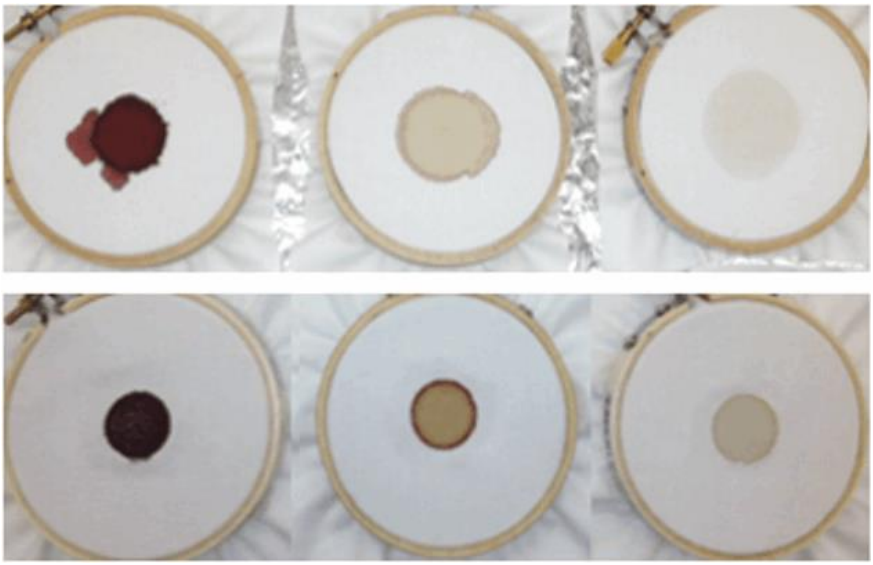

Figure 3.5. The bloodstains prepared in Figure 2 were allowed to dry overnight before applying one, 100 µL aliquot of luminol solution to each. Top: The result confirms that luminol applied dropwise does not spread in a reproducible manner, which ultimately forfeits knowledge concerning the true amount of blood and luminol reacting to produce a chemiluminescent response. Furthermore, this phenomenon renders the experiment non-reproducible. Bottom: Blood and luminol solutions occupy the same, reproducible area of cotton substrate.

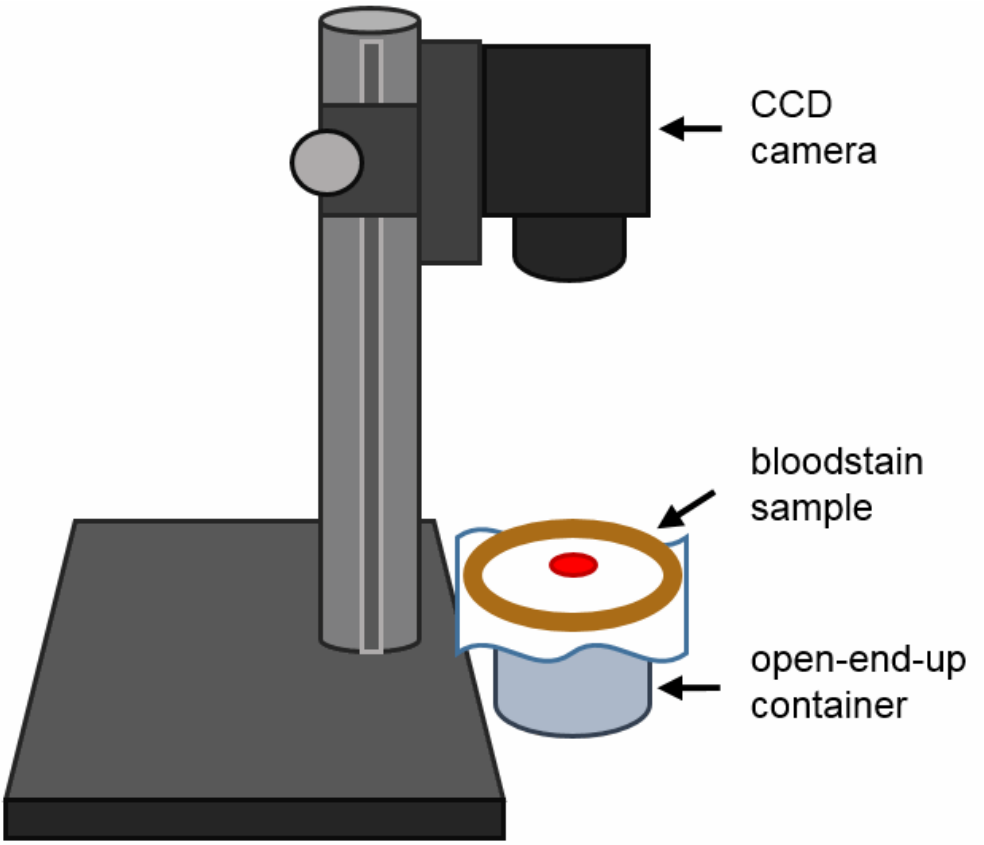

Figure 3.6. Experimental setup using a CCD camera for measuring chemiluminescent intensity response of luminol to blood.

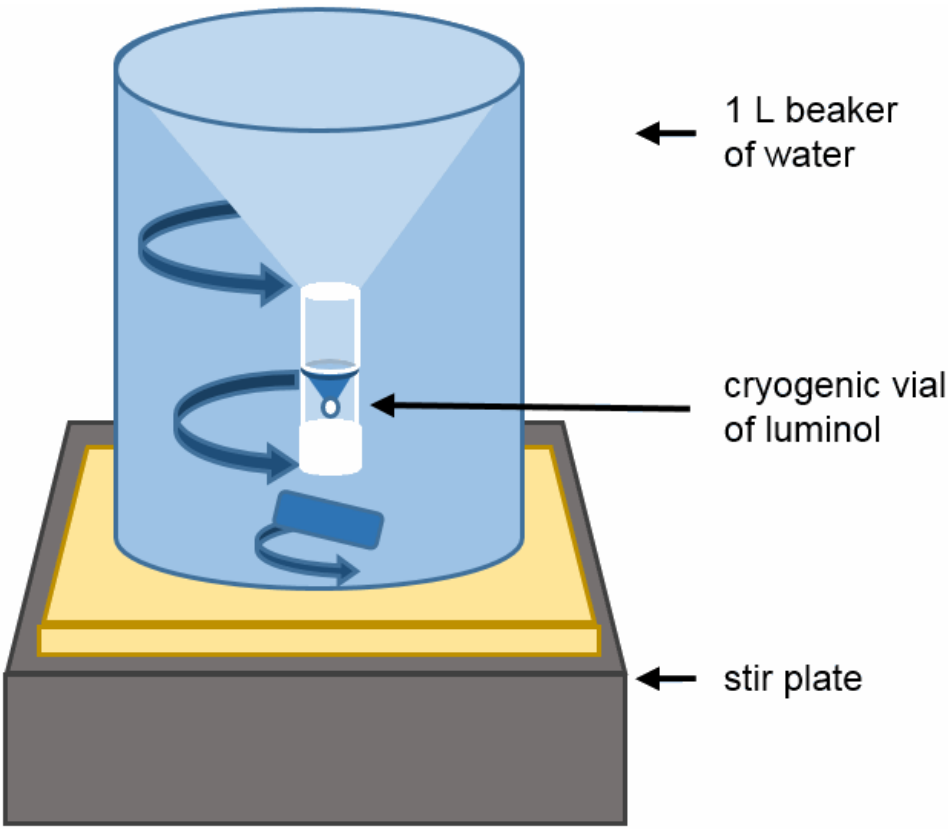

Figure 3.7. Controlled thawing environment for the luminol solution. A beaker was filled with tap water at room temperature and agitated with a stir bar at high enough speed to create a vortex to enable clear viewing of the melting point of luminol inside the cryogenic vial.

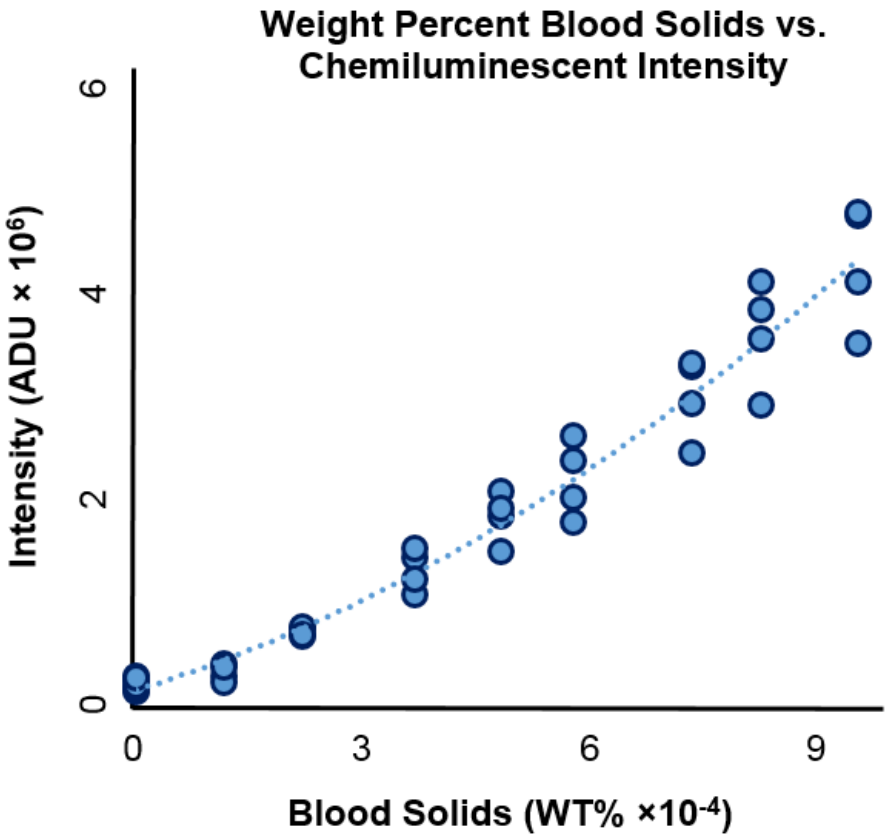

Figure 3.8. Calibration for the Bluestar® LD estimation on cotton substrate. Eight blanks and four replicates of each of eight bloodstain dilutions were tested using the same solution of 15 min. old Bluestar® solution. A Bluestar® LD for bloodstains of 9.16×10^{-5} WT% blood solids was determined. The fitted quadratic calibration equation is $y = 1.4906\times10^{5} + 4.8548\times10^{10}\, x + 1.0367\times10^{15}\, x^{2}$, with an R^2 of 0.959.

APPENDIX A – PLAYABILITY TEST

ABSTRACT

Playback of tapes on vintage tape players is the standard method for differentiating degraded from non-degraded tape. Tapes were categorized as non-playable if: (a) friction between the tape and player guides slowed or stopped the tape transport; (b) the tape produced squealing noises at any time; (c) the tape exhibited slow speed recovery, indicative of increased friction, between fast forward (FF) and rewind (RW) transitions; or (d) significant tape material sloughed onto player guides. Although tapes exhibit degradation effects during playback in different ways, tapes that showed any of these four behaviors were categorized as non-playable. Tapes were categorized as playable if they exhibited smooth and quiet playback in play, FF, and RW modes without sticking or shedding. Small amounts of edge shed and transfer of magnetic-layer material onto the guides was not considered indicative of non-playable tape. Videos which exhibit different playability status are present

METHOD

All tapes were played by a sound engineer at the MBRS using a Scully 280 tape player (Scully Recording Instruments, Bridgeport, CT) (Figure Appendix A-1).

Playability status was determined by attempting to pass the entire tape from one real to the next over six stationary guides (with the read and recording heads removed). All player guides were cleaned using cotton swabs and methanol in between tapes. Once the tape was threaded as illustrated by Figure S-1, the tape was assessed.

The playability test was carried out using the following protocol until the tape stopped the player or the whole tape was transferred to the take-up reel:

a. "PLAY" for 10 seconds.
b. STOP
c. "FF" then "RW" repeated 3 times.
d. STOP
e. Toggle between "FF" and "RW" to keep overall the tape movement moderate and continuous through about 30% of the tape.
f. STOP
g. Repeat steps a-f three times or until the tape stops the player.

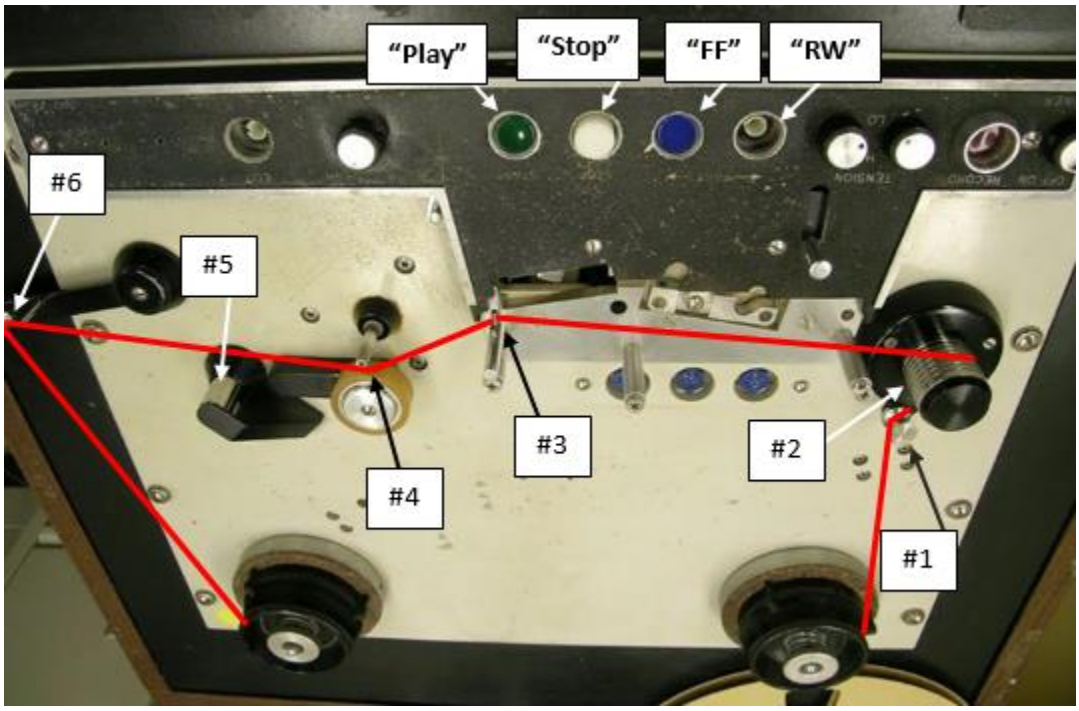

Figure A.1: The tape was threaded from the hub (not installed here) closest to guide 1 through the player, from guides 1 to 6, to the take-up hub (not installed here) closest to guide 6.

APPENDIX B – STAIN BARRIER APPLICATION

ABSTRACT

Determination of a LD depends heavily on the ability to create reproducible samples. However, the more dilute a blood solution, the further the solution spreads when applied to a substrate. This phenomenon complicates calculation and relationship of the mass of blood solids per amount of substrate and introduces an element of randomness from sample to sample. Stain barriers were implemented to limit the effect blood dilution has on spreadability and enhance sample reproducibility. The method of applying stain barriers to fabric substrate is presented here.

METHOD

First, the cotton substrate is placed in a 3″ embroidery hoop. Caps taken from screw thread vials (Fisherbrand®, CAT NO 03-339-21F, Pittsburgh, PA) are placed on each side of the fabric so that the open ends of the caps face each other with the fabric sample in between. The caps must be lined up precisely, so that a continuous seal is formed between the fabric inside the caps and outside the caps. A C-clamp (Grizzly Industrial®, Model: G8094, Springfield, MO) is used to securely hold the caps in place.

C-clamps are held in an upright position using clips here. However, any device able to hold the C-clamps upright without touching the cotton substrate is appropriate (Figure Appendix-B-1). The metal torque device on the C-clamps may be covered (here with clear tape and fabric) to prevent contamination. Using a cotton swab, one coat of PVC cement is applied liberally around the caps on each side of the cotton substrate. The fabric should be taut and level before allowing the glue to dry. We use cotton swabs as props in between the embroidery hoop and table top to level the substrate. The glue is allowed to dry for at least one hour before the C-clamp and caps are removed. More than one stain barrier can be applied to each substrate.

Figure B.1: The stain barrier application apparatus. One cap is place above and another is placed below (though not able to be seen here) the substrate. The caps are held in place by a C-clamp held upright with plastic clips. The PVC glue is placed around the caps on both sides of the fabric and is allowed to dry for one hour before the C-clamp and caps are removed.