Research and Training in CSAFE

Alicia Carriquiry
Iowa State University

November, 2020

CSAFE 2.0

- CSAFE 2.0 (Center for Statistics and Applications in Forensic Evidence) is a NIST Center of Excellence.
- Center established in 2015, recently renewed for an additional five-year period, until 2025.
- Consortium of six major academic partners:
  - Carnegie Mellon University (PI Robin Mejia)
  - Duke University (PI Brandon Garrett)
  - Iowa State University (PI Alicia Carriquiry, Director)
  - University of California Irvine (PI Hal Stern)
  - University of Virginia (PI Karen Kafadar).
  - West Virginia University (PI Keith Morris).
- Three affiliated institutions: Swarthmore College (Amanda Luby), University of Nebraska Lincoln (Susan VanderPlas), University of Pennsylvania (Maria Cuellar).

Our three-part mission

- Please visit www.forensicstats.org.
- **I. Research**
    - Focus is on development of statistical methods for pattern evidence and digital evidence.
    - Active research projects in various disciplines: footwear, firearm and tool marks, questioned documents, blood stain pattern and latent prints.
    - Cross-cutting areas include basic statistical research, research on implementation and communication strategies.
    - Significant effort spent on creating resources for researchers: databases, algorithms.
    - Everything in the public domain.

Our three-part mission

- **II. Training**
  - Audience: forensic practitioners, legal professionals, forensic science students (graduate and undergraduate), lay public.
  - In person and online, formats vary depending on topic.
- **III. Engagement and knowledge exchange**
  - A strong focus in CSAFE 2.0.
  - Aim is to translate research into practice.
  - Strategy is to increase communication, exchanges, collaboration with forensic professionals.

Overall goal

- The overall goal of CSAFE 2.0 is to **develop, test and validate** tools for forensic professionals to enable:
  - A more objective, science-based approach to the evaluation of evidence.
  - Quantification the degree of similarity between two items, e.g., two footwear prints.
  - Probabilistic statements about source of two items in evidence.

- Standard statistical tools are sometimes difficult to implement on non-standard data e.g., images.

- CSAFE researchers make use of a wide range of tools: statistical modeling, classification algorithms, more general machine learning methods.

Emphasis on engagement

- From the start, we have emphasized research with the potential to strengthen *the practice* of forensic science in the United States.
- Strong partnerships with forensic professionals and with the legal community are critical to our success:
  - You own the problems, understand the gaps.
  - The tools we develop must be tested in realistic environments.
- We have some funding to facilitate collaborations: examiners' time, materials, travel (eventually...).
- Training as a form of facilitating engagement.

The stakes are high

- Forensic evidence keeps gaining in importance, so valid, reliable tools are needed more than ever.
- Experienced examiners cannot be replaced by statistical and algorithmic tools, but tools can aid and complement their work.
- Recent court challenges limiting the strength of conclusions based on expert opinions serve as both motivation and guidance for the work ahead of us.

Specific opportunities

- **Statistical and Algorithmic Approaches to Matching Bullets and Cartridges**:
  - Opportunity for labs to help us evaluate and pilot new methods.
  - We will gladly scan bullets and cartridge cases and share both images and analyses.
- **Probabilistic Assessment of Handwriting**:
  - Participate in our data collection effort! Looking for at least 100 new participants.
  - Looking for database of signatures to test new methodology.
- **Blind Proficiency Testing**:
  - • Opportunity: Explore ways to implement blind testing by collaborating with other labs who have done this or are in the process.
- **Training and Education for Forensic Practitioners**:
  - Opportunity: Host training for lab or groups of labs.

Today's presentations + advertisement

- Highlights of three research projects in different areas.
  - Prof. Charless Fowlkes (Computer Vision, University of California Irvine): Footwear Impression Analysis.
  - Prof. Heike Hofmann (Statistics, Iowa State University): Algorithms for Firearms and Toolmarks.
  - Prof. Brandon Garrett (Law, Duke University): Issues on Implementation and Practice.
- Coming up:

   (Virtual) Workshop for firearm examiners organized by CSAFE.

  on November 30, December 1, 2020.
- To register, either go to www.forensicstats.org or (better) contact Harlie Jud at harliej@iastate.edu.
- Registration is free.

Footwear Impression Analysis

Presented by: Dr. Charless Fowlkes

forensicstats.org

CSAFE 1.0 Accomplishments

Overarching goal: *Devise the means to assess the strength of association between a crime scene print and a suspect's shoe and to gather the information that is needed to develop a score-based likelihood ratio framework for footwear examination.*

- Dataset Collection:
  - Longitudinal database of outsole impression data (160 pairs of shoes worn over 24 weeks, 30k images including test impressions and 2D/3D scans)
  - Shoe outsole imagery database (>200k retail tread photos of 80k products)

- Methods for feature-based comparison and matching:
  - Semi-automated score-based matching using random forests
  - Neural network-based feature matching and retrieval from crime scene prints (state-of-the-art performance on public benchmarks)
  - Detecting and cataloging frequencies of tread element types

- Probabilistic models describing non-uniform spatial distribution of randomly acquired characteristics (RACs)

CSAFE 2.0 Research Area Objectives

Overarching goal: *Devise the means to assess the strength of association between a crime scene print and a suspect's shoe and to gather the information that is needed to develop a score-based likelihood ratio framework for footwear examination.*

Objectives:

- Existing models to understand the spatial distribution of RACs rely on simplifying independence assumptions. We aim to develop and validate models that are more realistic for RACS.

- A limitation in footwear analysis is the absence of reliable, publicly available databases that characterize the distribution of footwear characteristics and patterns.

  - Collect tread product imagery available online to characterize diversity of tread patterns

  - Develop an instrument that can be deployed in public areas to passively collect images of outsoles of shoes. These data can enable estimation of the frequency of footwear patterns in a given region and facilitate the transition to a likelihood ratio or Bayes Factor approach to the evaluation of footwear evidence.

  - Assemble a database of representative simulated crime-scene evidence with known ground-truth

- Extend statistical approaches to feature-based matching of crime scene prints developed during CSAFE 1.0 to be more robust so that they are viable with partial and obscured evidence images.

- Develop evidence quality metrics for footwear impression images that are predictive of matching uncertainty (for either examiners or algorithms) to aid interpretation of footwear evidence.

CSAFE 2.0 Footwear Projects

FW I - Occurrence and Persistence of Accidental Mark Locations On a Shoe Sole
- Lead PI: Hal Stern, UC Irvine

FW II - Statistical Models for the Generation and Interpretation of Footwear Impression Evidence
- Lead PI: Charless Fowlkes, UC Irvine

FW III - Characterization of Footwear in Local Populations
- Lead PI: Susan Vanderplas, UNL

FW IV - Statistical and Algorithmic Approaches to Shoeprint Analysis
- Lead PI: Alicia Carriquiry, ISU, Jacqueline Speir, WVU

***** –Building a Public Dataset of Simulated Crime Scene Impression Evidence**
- All team PIs (cross-cutting)

Ongoing Collaborators and Partners:

Hari Iyer (NIST), Steven Lund (NIST), Martin Hermann (NIST), Sarena Wiesner (Israeli Police), Aviad Levi (Israel Police), Yoram Yekutieli (Hadassah Academic College), Micha Mandel (Hebrew University)
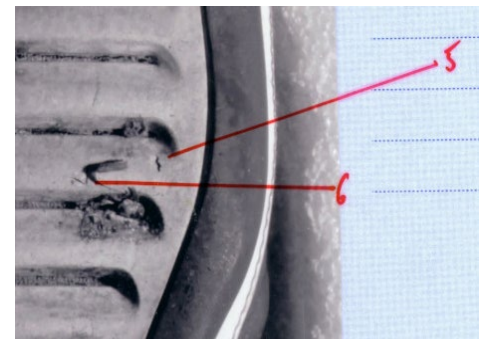
CSAFE 2.0 Footwear

Occurrence and Persistence of Accidental Mark Locations On a Shoe Sole

Activities:

- Study the reliability of RAC identification, both between-examiner (reproducibility) and within-examiner (repeatability).
  - Methods developed should address location, type, size and shape.

- Collect data (in concert with other FW projects) and use data to understand differences between observed RACs on a shoe and on a crime scene print. Plan to address the following questions:
  - What is the probability of observing a RAC at a crime scene, and how does it depend on the shape, type, size, location and other properties of the physical RAC?
  - How does the probability of observing a RAC depend on the substance and the quality of the trace?

- Build a statistical model to estimate the intensity function of RAC location and account for dependency among RACs and variabilities among individuals

Impact:

- Contribute to the understanding of the scientific strength of footwear comparisons

- Assist in accurate determination of the value of correlation found between the crime scene print and the shoe
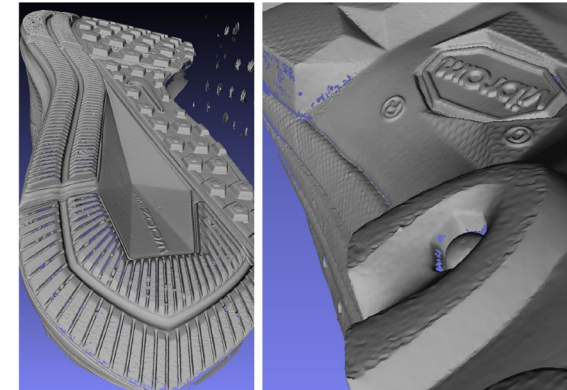
CSAFE 2.0 Footwear

Statistical Models for the Generation and Interpretation of Footwear Impression Evidence

Activities:

- Develop statistical matching models for reasoning about partial or obscured prints
  - Computational feature matching to provide calibrated match scores of class characteristics from partial prints
  - Generate a large-scale library of tread patterns derived from product imagery to analyze diversity of class characteristics

- Methods for relating three-dimensional (3D) tread shape to impression evidence
  - Models for estimating contact surface and impression likelihood from high-resolution 3D scans
  - Computer vision techniques for estimating contact surface from commercial tread imagery

Impact:

- Statistical basis for better understanding the reliability of impression evidence in determination of class characteristics

- Assistive tools for automatically searching databases of tread patterns and calibrating reliability

CSAFE 2.0 Footwear

Characterization of Footwear in Local Populations

Activities:

- Collect footwear data from populations interacting with law enforcement and corresponding general populations.
- Assess class characteristic distribution and variation over time and between populations.

Impact:

- Develop methods to support random match probability calculations for footwear class characteristics.
- Establish differences in reference populations.
- Understand variability in class characteristic frequency related to time and location

CSAFE 2.0 Footwear

Statistical and Algorithmic Approaches to Shoeprint Analysis

Activities:

- Continued development of CSAFE 1.0 semi-automated score-based matching techniques
- Propose and evaluate automated image descriptors (noise, texture, frequency, totality, etc.).
- Obtain crowdsourced expert-opinions of image quality.
- Obtain numerical/objective metrics of source association for pairwise comparisons.
- Regress expert-assessed quality and algorithmic similarity with quality metrics in order to assess goodness-of-fit.
- Create predictive models of image quality for novel imagery.

Impact:

This research will inform quality, value and sufficiency predictors that can be used to qualify uncertainty and error estimation in source association conclusions.
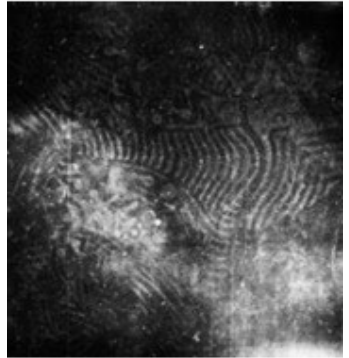
CSAFE 2.0 Footwear

Assembling an Open Dataset of Simulated Crime Scene Impression Evidence

Activities:

- Assemble a dataset consisting of simulated crime scene evidence with *known-source* test impressions

  - Covering a diverse range of outsole designs, sizes, wear patterns and accidentals

  - Spanning a variety of impression types (dry/wet, chemical enhancement, lifting techniques) and substrates (wood, tile, paper)

  - Generated by a range of activities (e.g., kicking a door, jumping onto a counter, walking, running)

- Prioritize collection of data which is representative of examiner caseloads

  - Carrying out survey of practitioners to determine what types of evidence are most common

Impact:

This effort will provide essential data in support of a wide range of research including modeling impression formation, variability in impressions, repeatability of RAC appearance, evidence quality assessment,

Resources and Needs

Looking for collaborators:

- Labs and examiners willing to participate in surveys of how frequently different types of evidence, collection techniques and brands of shoes are encountered in typical casework

- Exploring fine grained proficiency testing to understand repeatability / reliability of marking individual RACs

- Correlating retail footwear sales data with distribution of footwear worn by local populations

CSAFE Footwear Resources:

- Periodic Webinars + Insights

- Datasets and other resources
  https://forensicstats.org/

Firearms and Toolmark Analysis

Presented by: Dr. Heike Hofmann

Research Area Objectives

Currently established comparison standard:

AFTE Theory of identification

1. examine class characteristics
2. use microscopic analysis to assess detailed features

Identified Problems:

1. establishing error rates of identification process
2. subclass characteristics (determined by proficiency tests in Europe) are a key risk factor for false identifications.

CSAFE 1.0
Accomplishments
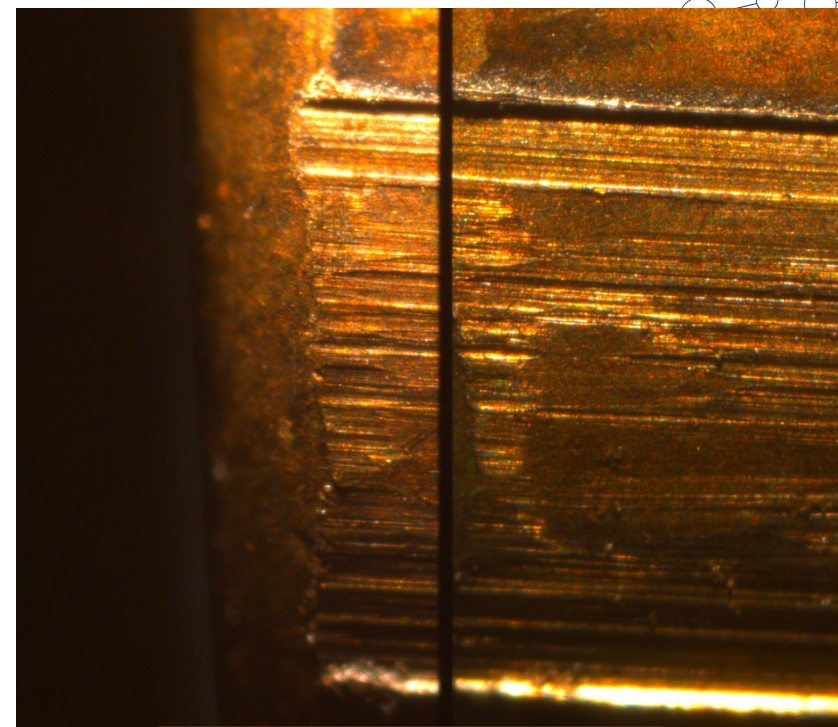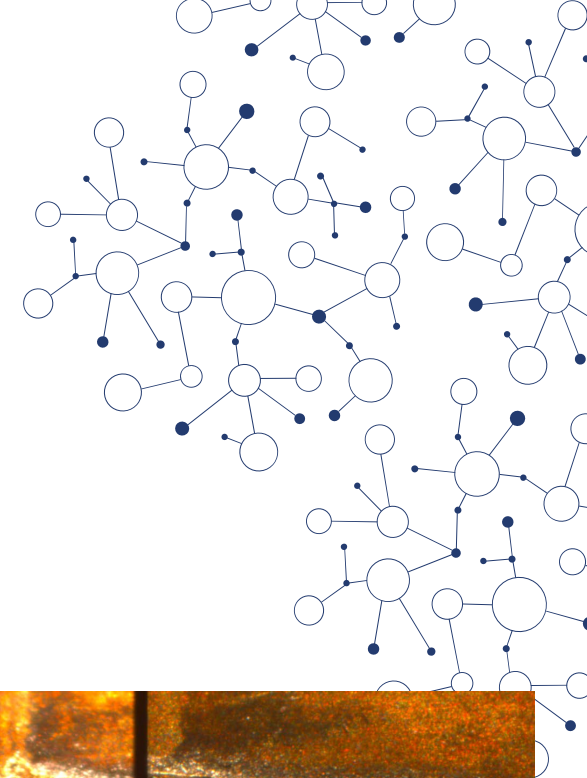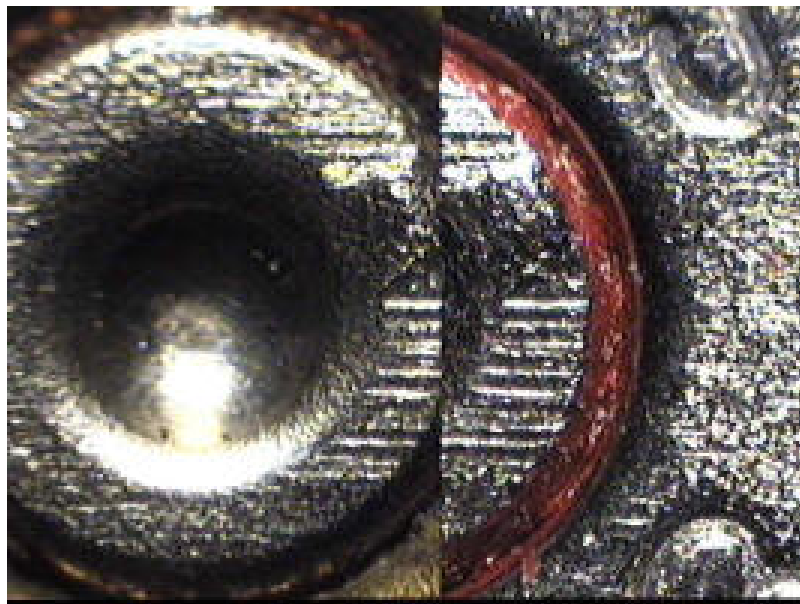
CSAFE 1.0 Accomplishments

1. Two automated **matching algorithms**:
   bullets (Hare et al., 2017a; Hare et al., 2017b)
   cartridge cases (Tai and Eddy, 2017)

2. **Open source** algorithms: *cartridge3D, x3ptools, bulletxtrctr*

3. **Open data**: 3d topographic high-resolution scans of bullet lands (~25,000) and cartridge cases (~2,000), mostly uploaded to the NIST Ballistics Toolmark Research Database Evaluation (NBTRD).

CSAFE 2.0 Objectives
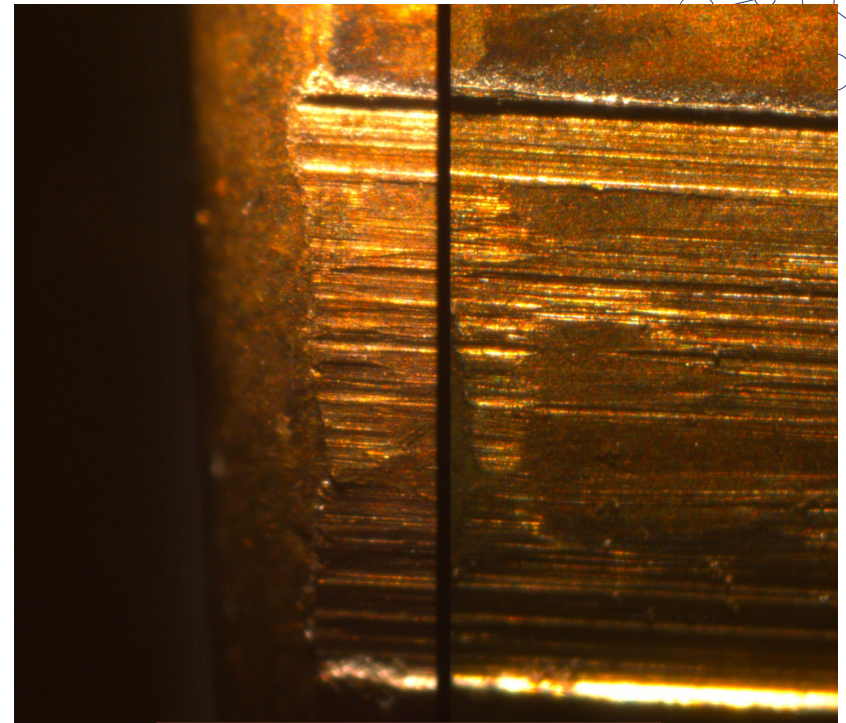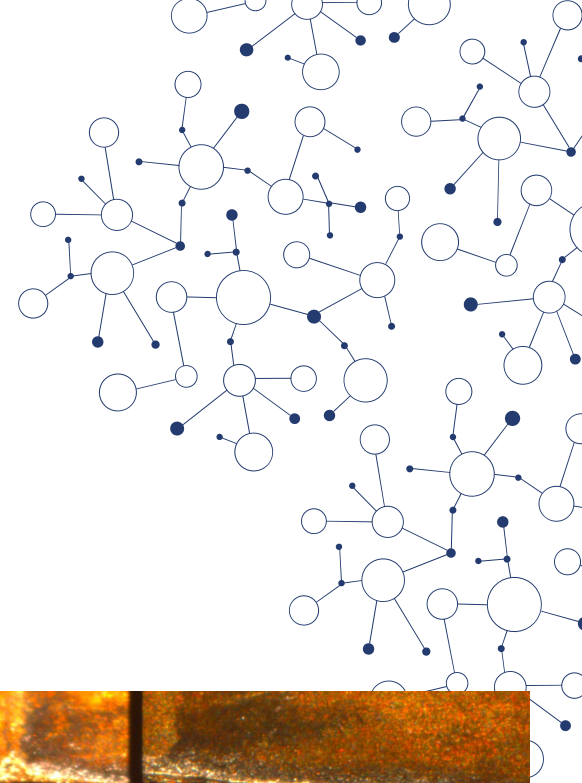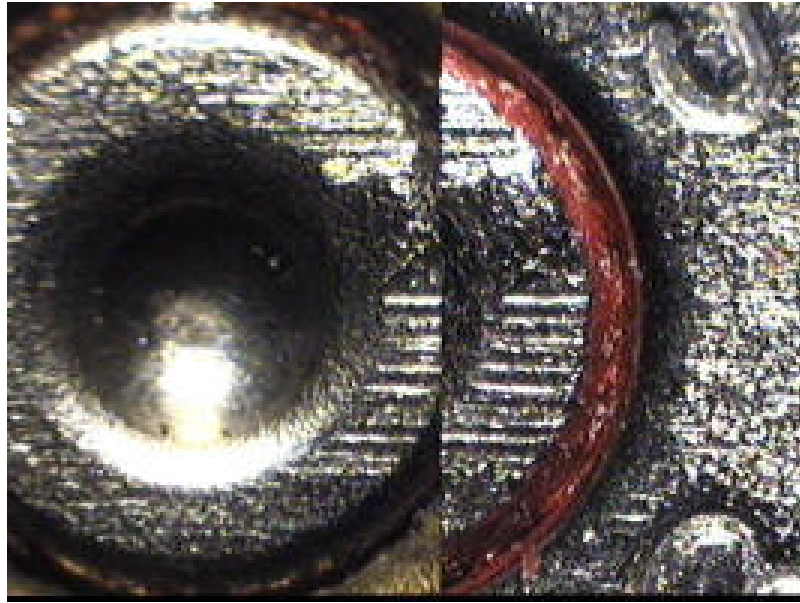
# CSAFE 2.0 Projects and Lead Investigators

F&T I- Statistical and Algorithmic Approaches to Matching Bullets and Cartridges
Lead PI: Heike Hofmann, ISU

F&T II- Subclass Characterization and Analysis of Firearms
Lead PI: Keith Morris, WVU

F&T IV- Evaluating Foundational Validity of Toolmark Analysis
Lead PI: Maria Cuellar, UPenn

F&T I Statistical and Algorithmic Approaches to Matching Bullets and Cartridge Cases
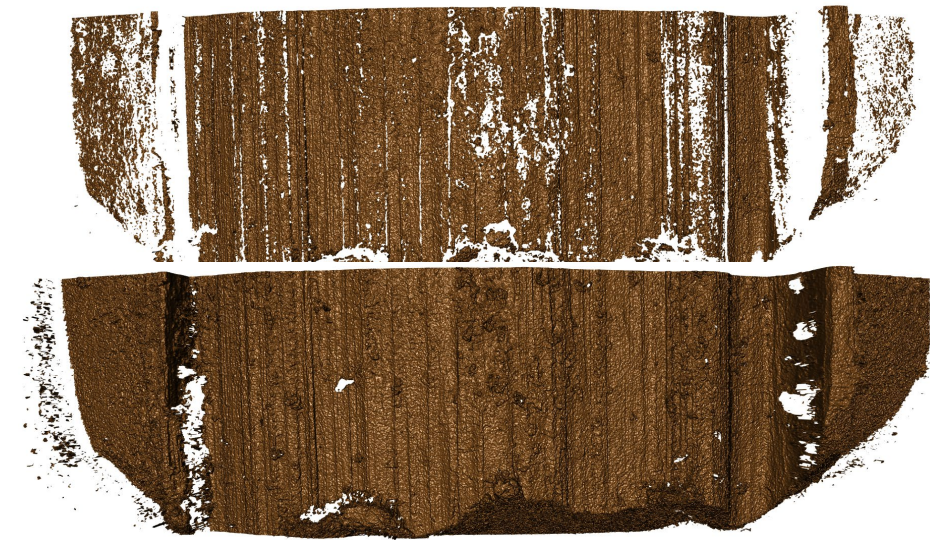
Proposed Activities:

- Expand and refine matching algorithm: nontraditional rifling, new features based on image

- Quantify factors affecting matching performance: combination of firearm/ammunition, quantitatively assess quality of scans.

- Work with firearms examiners and crime labs to extend use of matching algorithms to labs

Potential Impact:

- Providing objective quantitative assessments that examiners can use during testimony

- Providing empirical support for the validity if firearms and toolmarks evaluation through objective algorithmic assessments

- Working with examiners to develop community confidence and trust in algorithmic results

F&T II Subclass characterization and analysis of firearms

Proposed Activities:

- Reference collection of Contender G2 breech faces
- Collection of reference collections from five forensic laboratories
- Characterization of the manufacturing processes and breech faces
- Collaborate with firearm examiners to identify areas of subclass on all breech faces and test fires
- Automated comparison using NIST congruent matching cells (CMC) algorithm with and without subclass characteristics present
- Creation of subclass markup GUI

Potential Impact:

- Test sets can be created from test fires
  (both with digital scans and double-castings)
- Examiner accuracy testing of identifying subclass characteristics
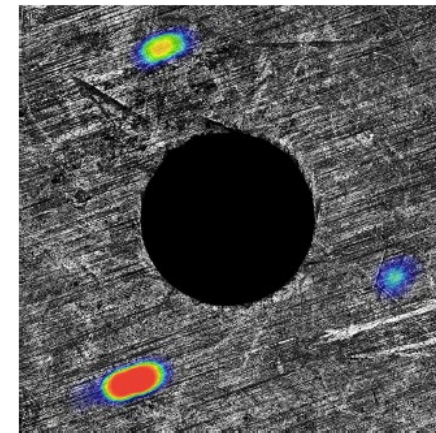- Performance of NIST CMC algorithm with subclass present

Figure 7: Example heatmap of subclass characteristics. Refer to the electronic version for color.

F&T IV Evaluating Foundational Validity of Toolmark Analysis

Proposed Activities:

1. **Database:** Create a database of high-quality toolmark images, both 2D and 3D, using a factorial design, based on NBIDE firearms database.

2. **Algorithm:** Develop an algorithm to determine a score-based likelihood ratio.

3. **Validation:** Validate algorithm by testing its external validity.

Potential Impact:

- **Start with simplest case to make progress in difficult field** of toolmark analysis, which has many types of tools and degrees of freedom.

- **Develop standard statistical methods** for the analysis and comparison of toolmarks.

- **Expand the capacity** of federal, state, and local labs to deal with toolmark analysis.

**Create new database of toolmarks (start in 2D, then 3D)**

NBIDE:
- Firearm brand
- Firearm # (of same brand)
- Ammunition brand
- Iteration per firearm/ammo combination
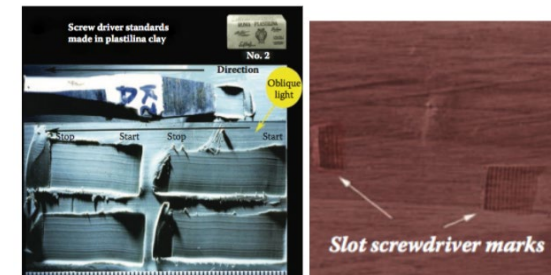
Iteration #1...

Toolmarks:
- Screwdriver brand
- Screwdriver # (of same brand)
- Angle, surface material (soft to hard), striation vs. imprint, etc.
- Iterations per angle/surface/etc. combination
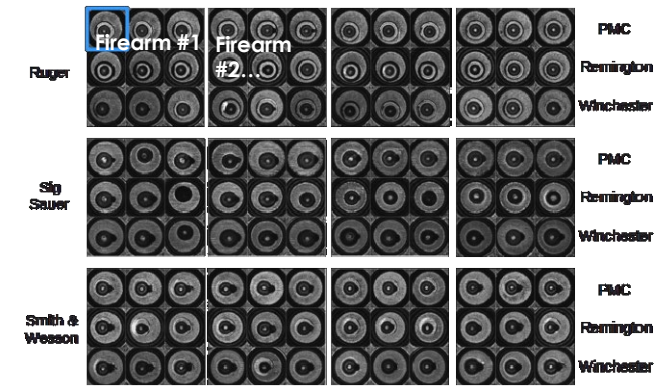
Many degrees of freedom in screwdriver mark
Baiker et al. (2016).
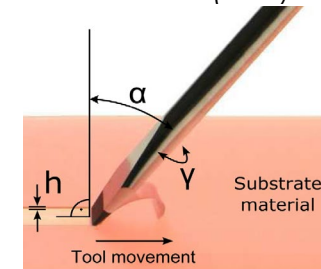
Screwdriver striation (L) and impression marks (R).
Petraco (2011).

Your help is needed !!!

https://forensicstats.org/

hofmann@iastate.edu
alicia@iastate.edu
srenfro@iastate.edu

- Forensic examiners community:
  - Standard operating procedures for assessing firearm and toolmark evidence of labs
  - Test fires of (some of) the reference collections – format needs to be determined, but ideally we would like 4 test fires for each firearm/ammunition combo
  - AFTE studies (past and on-going): we would be excited to get materials for 3d imaging!

- Forensic analysts:
  - Help us in running (proprietary) algorithms on publicly available data and make results available
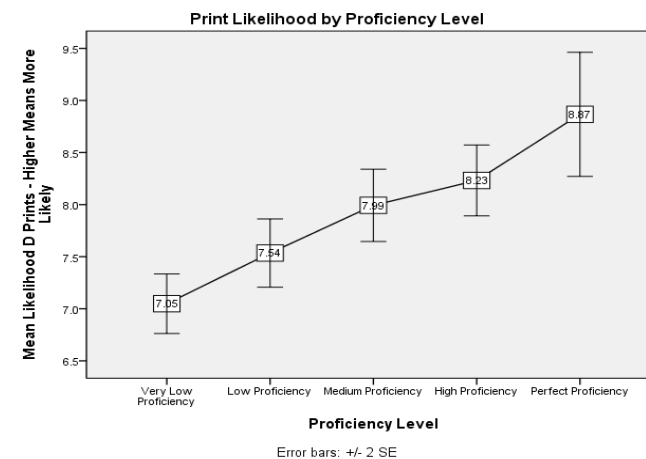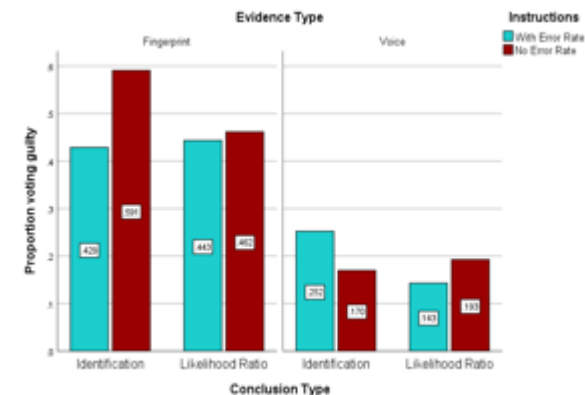
CSAFE 1.0 Accomplishments

Project Title: Project U - Research on Lawyers, Jurors, and the Evaluation of Forensic Evidence

Major Accomplishments:

- Thompson, W.C.; Scurich, N. "How Cross-Examination on Subjectivity and Bias Affects Jurors' Evaluations of Forensic Science Evidence," 64 J. Forensic Sciences (2019), doi: 10.1111/1556-4029.14031.

- Thompson, W.C.; Scurich, N. "When Does Absence of Evidence Constitute Evidence of Absence," For. Sci International 291 (2018).

- Thompson, W.C.; Scurich, N.; Dioso-Villa, R., Velasquez, B. "Evaluating Negative Forensic Evidence: When Do Jurors Treat Absence of Evidence as Evidence of Absence," 14 J. Empirical Leg. Stud. 569 (2017).

- Garrett, Brandon L.; Mitchell, Gregory, *"The Impact of Proficiency Testing Information and Error Aversions on the Weight Given to Fingerprint Evidence,"* 37 Behavioral Science and Law 1 (2019), doi: 10.1002/bsl.2402

- Mitchell, Gregory, Scurich, Nicholas, Garrett, B.L., Comparing Categorical and Probabilistic Fingerprint Evidence, J. Forensic Sciences (2018), doi: 10.1111/1556-4029.137 97.

- Garrett, Brandon L.; Mitchell, Gregory, *"The Proficiency of Experts,"* 163 University of Pennsylvania Law Review, Vol. 66, 2018.

- Garrett, Brandon L, Grady, Rebecca, & Crozier, William, *Likelihood Ratios, Error Rates, and Jury Evaluation of Forensic Evidence*, Journal of Forensic Sciences (2020)

- Garrett, Brandon L.; Mitchell, Gregory, Forensics and Fallibility: Comparing the Views of Lawyers and Jurors, 119 W.Va. L. Rev. 621 (2016)

Impact:

Unpacked how jurors respond to information about limitations of forensic techniques, including error rates, likelihood rations, and through proficiency, as well as work on quantitative expressions of conclusions. Informed ALI Principles, training for lawyers, and standards discussions, as well as judges. Publications for forensic practitioners, judges, and in law reviews.

CSAFE 1.0 Accomplishments

Project Title: Analysis of Forensic Testimony and Reports

Major Accomplishments:

- Completed collection and analysis of baseline of probabilistic reporting in forensic reports in 4 in-scope disciplines

- Completed survey of friction ridge examiners on attitudes toward probabilistic reporting with more than 300 responses in cooperation with the Defense Forensic Science Center.

- Completed survey of crime laboratory directors on probabilistic reporting.

- Began project of coding the cases involving forensic science contained in the National Registry of Exonerations, the nation's authoritative data repository of exonerations in cooperation with the Innocence Project.

Impact:

- Simon A. Cole & Matt Barno, "Probabilistic Reporting in Criminal Cases in the United States: A Baseline Study," revise and resubmit, *Science & Justice*
    - Presented at American Academy of Forensic Science

- Simon A. Cole, Henry Swofford & Valerie King, "'Mt. Everest--We Are Going to Lose Many': A Survey of Fingerprint Examiners Attitudes toward Probabilistic Reporting," in preparation.

- Simon A. Cole, "Individual and Collective Identification in Contemporary Forensics," *BioSocieties* (2018), https://doi.org/10.1057/s41292-018-0142-z.

- Simon A. Cole, "A Discouraging Omen: A Critical Evaluation of the Approved Uniform Language for Testimony and Reports for the Forensic Latent Print Discipline," *Georgia State University Law Review*, Volume 34, Number 4 (Summer 2018), pp. 1103-1128.

- Simon A. Cole & Barry C. Scheck, "Fingerprints and Miscarriages of Justice: "Other" Types of Error and a Post-Conviction Right to Database Searching," *Albany Law Review*, Volume 81, Number 3 (2017/2018), pp. 807-850.

CSAFE 1.0 Accomplishments

Title: Evaluating Lay Perceptions of Forensic Evidence (Project I)

Major Accomplishments:

Multiple peer-reviewed publications

questionnaires and jury simulation studies

participants recruited online and from jury pools

Commentaries in law reviews and professional publications for judges

This work has influenced discussions (OSAC and elsewhere) of how best to present results in reports and testimony

Research Area Objectives

As new probabilistic and statistical methods are developed they must be implemented by the forensic community.

CSAFE is focused on research to facilitate the implementation of methods and to identify best practices for their use.

Relevant research topics and goals:
Best ways to communicate statistical conclusions (both verbally and graphically),
Understanding the barriers to widespread implementation
Best practices for forensic practitioners, lawyers and judges.

CSAFE 2.0 Projects and Lead Investigators

IMPL I- Evaluating Lay Perceptions of Forensic Evidence
Lead PI: Brandon Garrett (Duke), Bill Thompson (UCI)

IMPL III- Understanding the Barriers to Accepting Probabilistic Methods
Lead PI: Simon Cole (UCI)

CSAFE 2.0

Garrett, Scurich, Crozier, 2020

Project Title: Project IMPL I - Evaluating

Lay Perceptions of Forensic Evidence

Proposed Activities:

1. **Efforts to better understand, as a foundational matter, how laypeople evaluate forensics**
    - **More realistic, ecologically valid designs. Up-to-date terminology. Deliberation studies.**
    •**Just published:** Garrett, Brandon L, Scurich, Nicholas, & Crozier, William, Mock Juror Evaluation of Firearms Evidence, 41 Law & Hum. Behav. 412 (2010)
    •Garrett, Brandon L, Kukucka, Jeff, & Crozier, William, Juror Appraisals of Forensic Evidence: Effects of Blind Proficiency and Cross-Examination, 315 Forensic Science International (2020)
    •Garrett, Brandon L.; Mitchell, Gregory, Creating Reasonable Doubt in Fingerprint Identification Cases: Substantive and Methodological Rebuttals by Defense Experts (under submission)
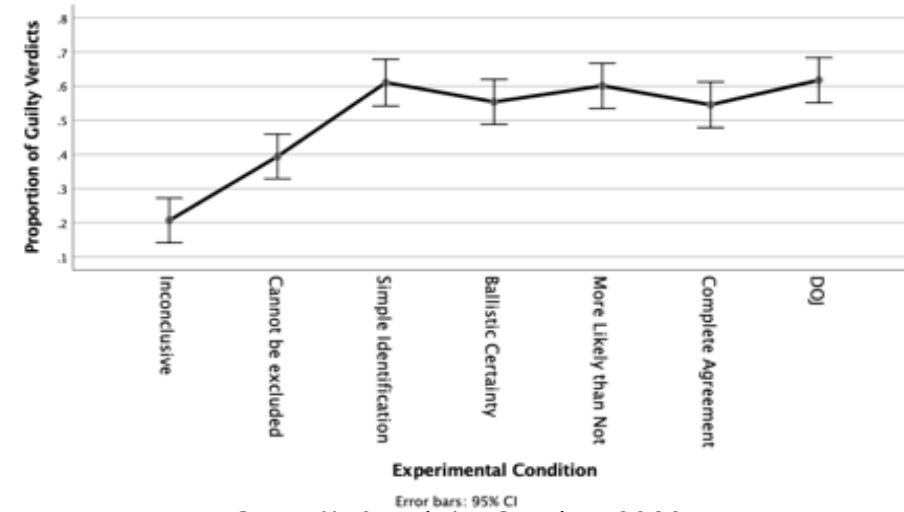
2. Testing new interventions, such as new language from OSAC, jury instructions, role of lawyers
    •Thompson, assessing lay reactions to new footwear reporting language

3. Develop model recommendations for how to express forensic results In testimony, to inform standards, judicial rulings,legal approaches towards litigating forensics

Potential Impact:

Developing empirical base for standard reporting language, judicial approaches, and training for lawyers

CSAFE 2.0

Project Title: Understanding the Barriers to Accepting Probabilistic Methods

Proposed Activities:

- Undertake sociological study of forensic statistics using social scientific methods, such as interviews, participant-observation, and ethnography, and history of science, such analyses of scholarly debates conducted through published literature.

- Replicate survey of practitioner attitudes toward probabilistic reporting for other in-scope pattern disciplines.

Potential Impact:

- Facilitate the adoption of statistical applications by enabling a better understanding of the organizational cultures of the forensic service providers that will have to adopt them.

- Enhance non-statisticians' understanding of the statistics discipline and the nature of the applications it is seeking to implement with a sociological account of the discipline.

CSAFE 2.0

Project Title: Evaluating Lay Perceptions of Forensic Evidence

Proposed Activities:

More elaborate and realistic studies of how communication of forensic science findings is affected by:

Characteristics of the messenger and message (e.g., nature of testimony; lawyers' arguments; graphics)

Characteristics of audience (e.g., education; numeracy)

Greater involvement practitioners in identification of research questions and preparation of experimental materials

Potential Impact: This work will be directly relevant to ongoing efforts to standardize and improve reporting practices in pattern matching disciplines.

Resources and Needs

- More outreach to lawyers, judges, forensic practitioners
- Please participate in surveys and experiments.
- Please consider starring in them!
- We look forward to your suggestions and feedback!