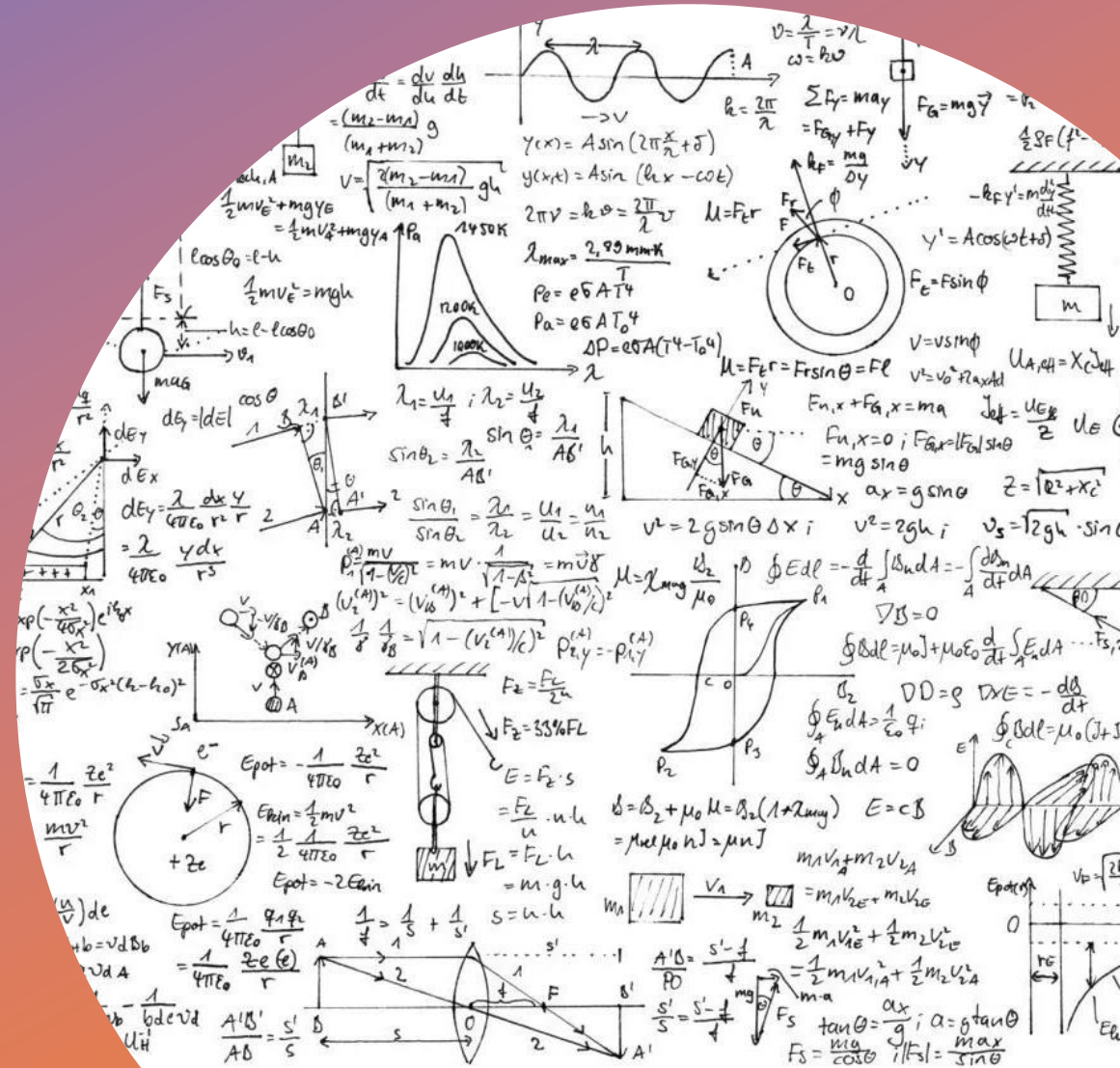# Applying Algorithms to Forensic Science

Jeff Salyards, PhD, MFS
Forensic Research Scientist IV

Center for Statistics and Applications in
Forensic Evidence (CSAFE)
Iowa State University

Brogan, Jacob. "What's the Deal With Algorithms? Your 101 guide to the computer codes that are shaping the ways we live. Slate Feb 02, 2016. https://slate.com/technology/2016/02/whats-the-deal-with-algorithms.html

Can I level with you? I'm not always sure I know what people are talking about when they say *algorithm*?

You're not alone: Honestly, I haven't always been sure what I meant when I said it either. But here's the absolute simplest definition: *An algorithm is a set of guidelines that describe how to perform a task.*

THE *NEW YORK TIMES* BESTSELLER

THE **CHECKLIST** MANIFESTO

HOW TO GET THINGS RIGHT

PICADOR

ATUL GAWANDE

BESTSELLING AUTHOR OF *BETTER* AND *COMPLICATIONS*

Do they make everything better?

Cooking
Aviation
Medicine
Forensic Science

Astronaut Neil Armstrong – Apollo 11
July 16, 1969
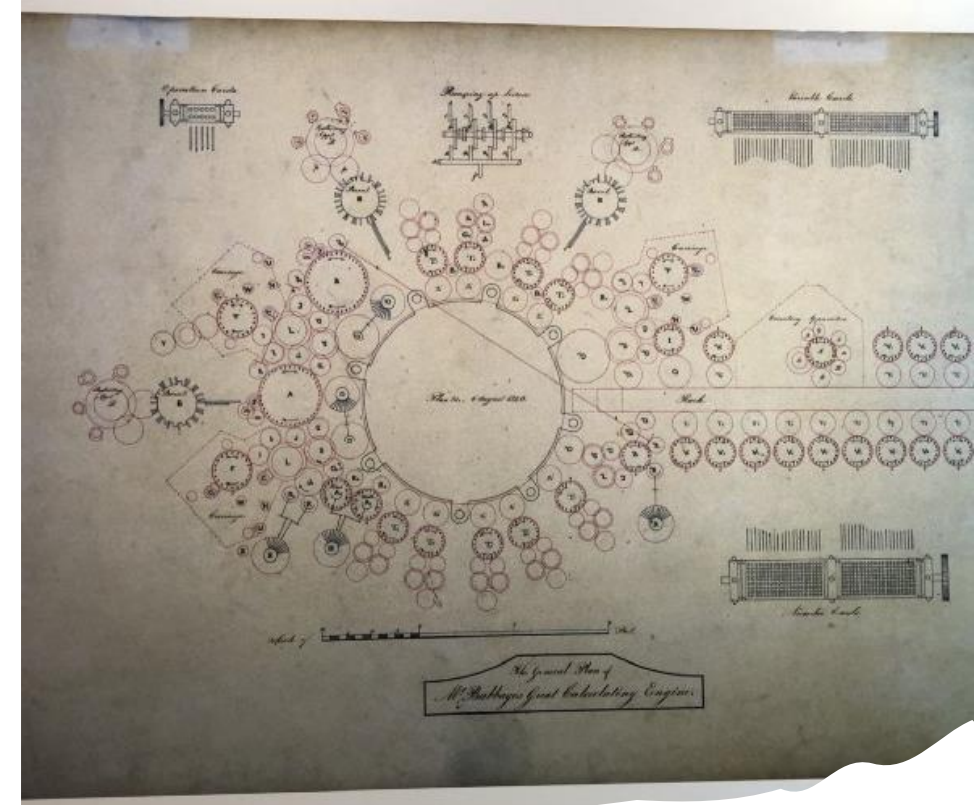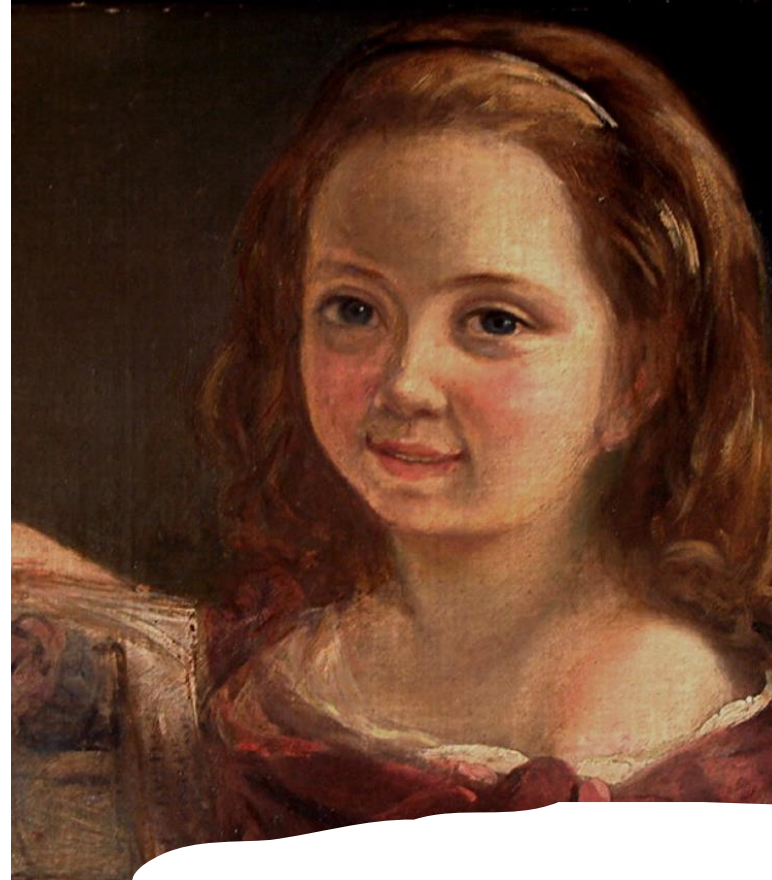
$$B^? = 1/2$$
$$B^2 = 1/6$$
$$B^4 = -1/30$$
$$B^6 = 1/42$$
$$B^8 = -1/30$$
$$B^{10} = 5/66$$
$$B^{12} = -691/2730$$
$$B^{14} = 7/6$$

Ada Lovelace, Charles Babbage, the Analytical Engine, Bernoulli numbers
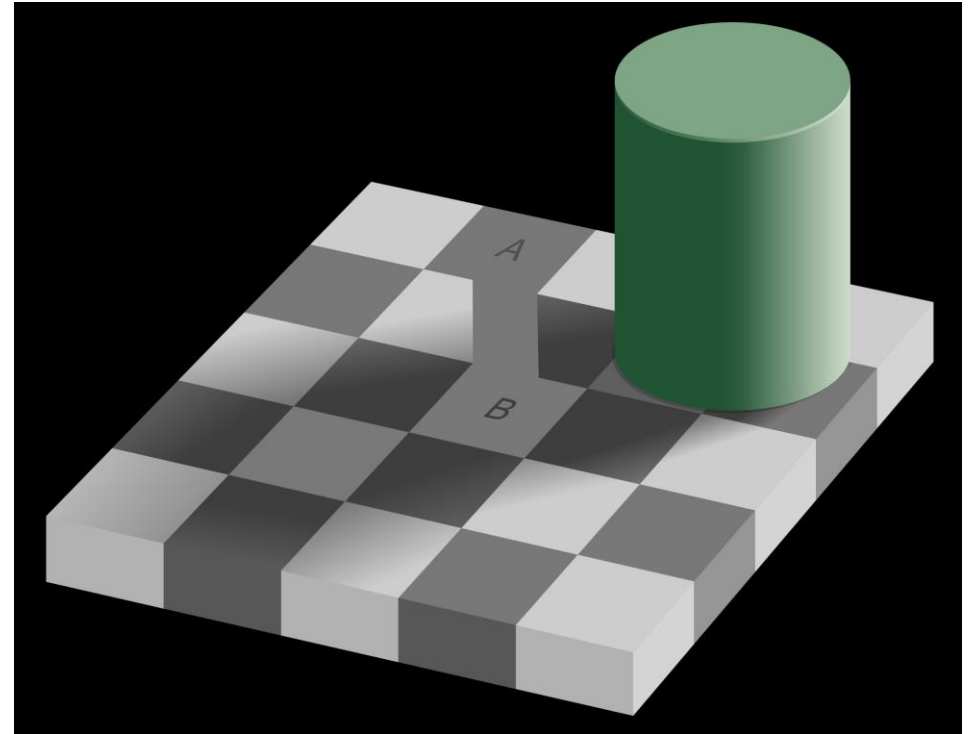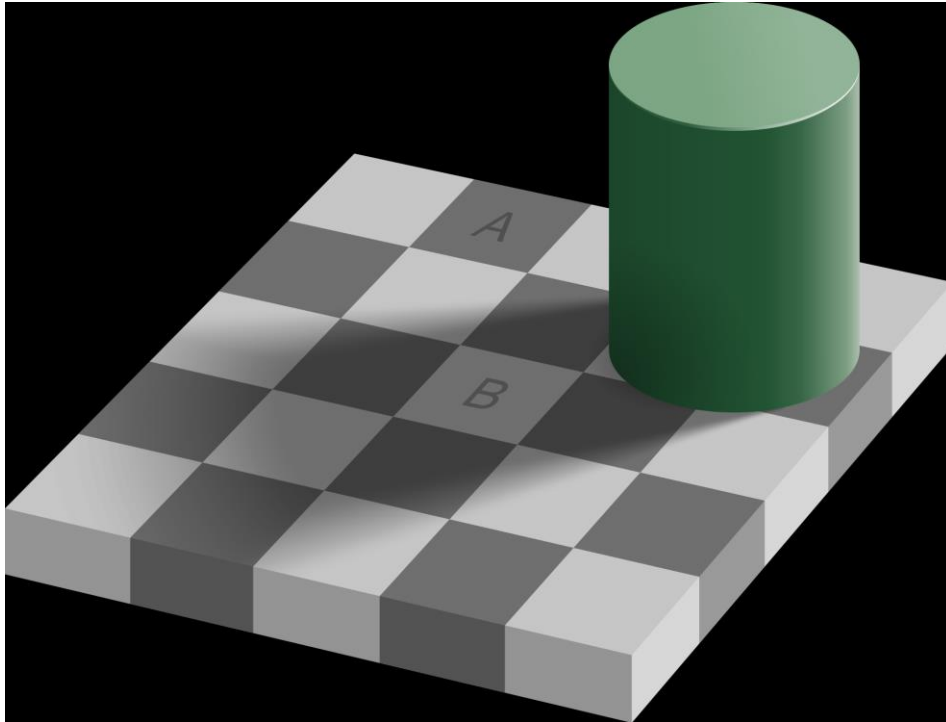
Our perception and intuition ... exclusionary difference?

Our perception and intuition ... exclusionary difference

David Novick, @NovickProf

David Novick, @NovickProf
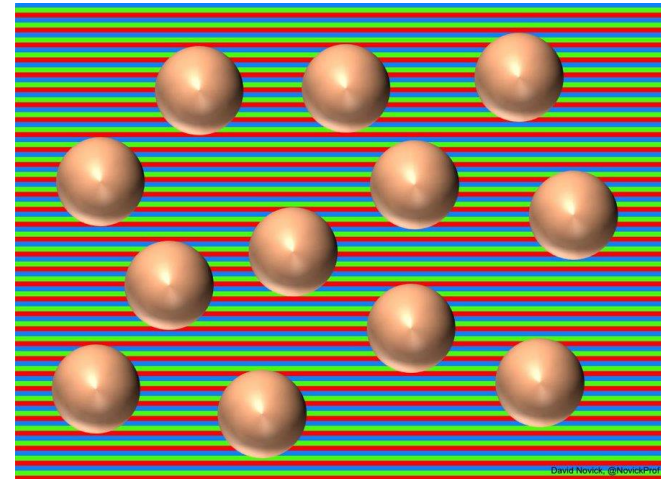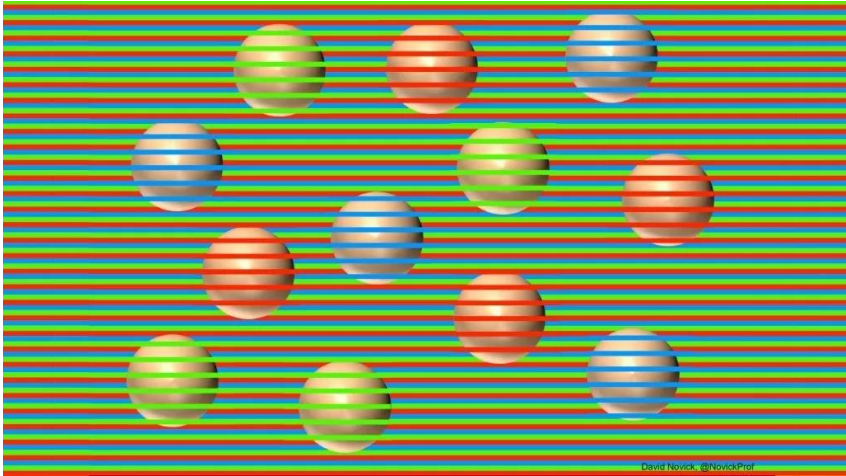
Our perception and intuition … fingerprint minutia

Our perception and intuition ... land and groove impressions

Professor Plum in the Library with the Gun? Knife? Candlestick?

Project BPA.I – Statistical Methods for Bloodstain Pattern Analysis

Accuracy and reproducibility of conclusions by forensic bloodstain pattern analysts

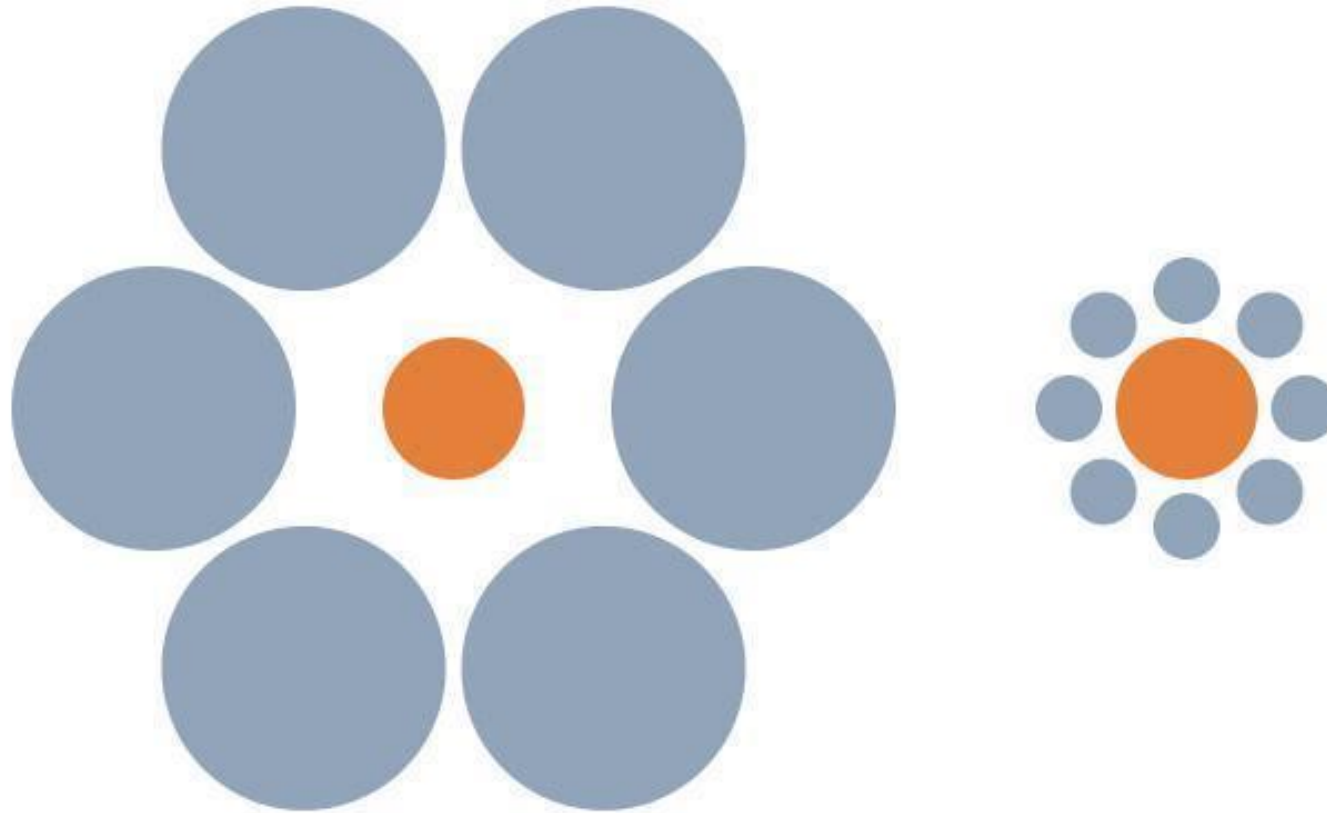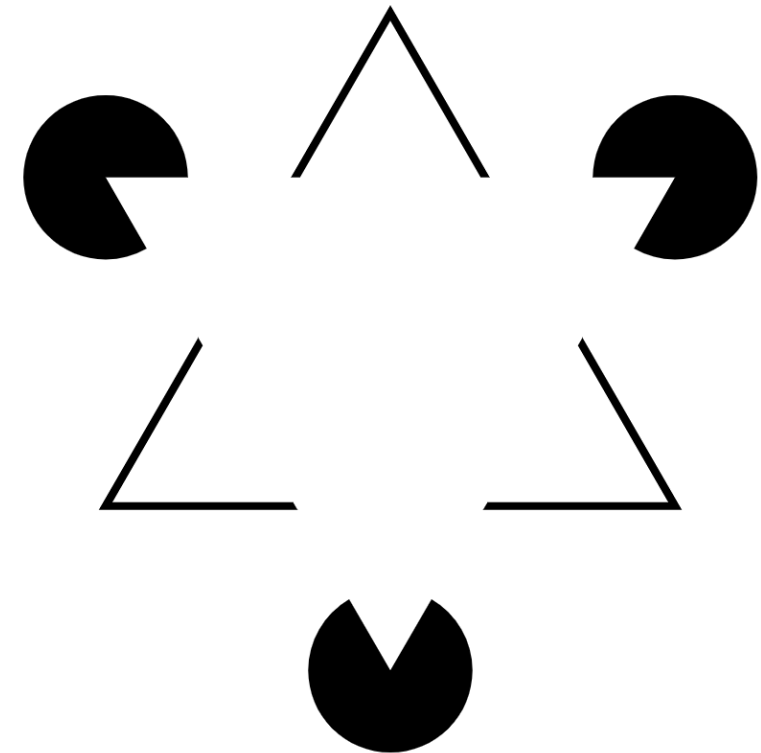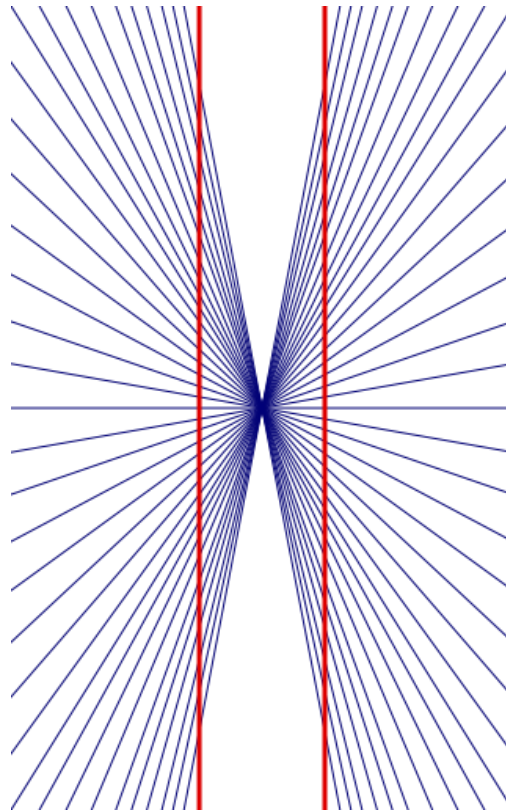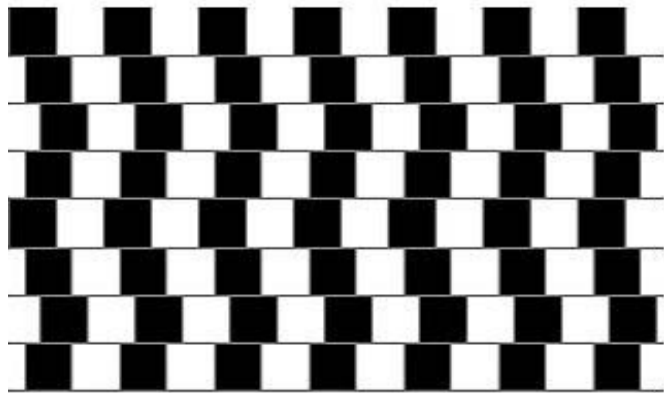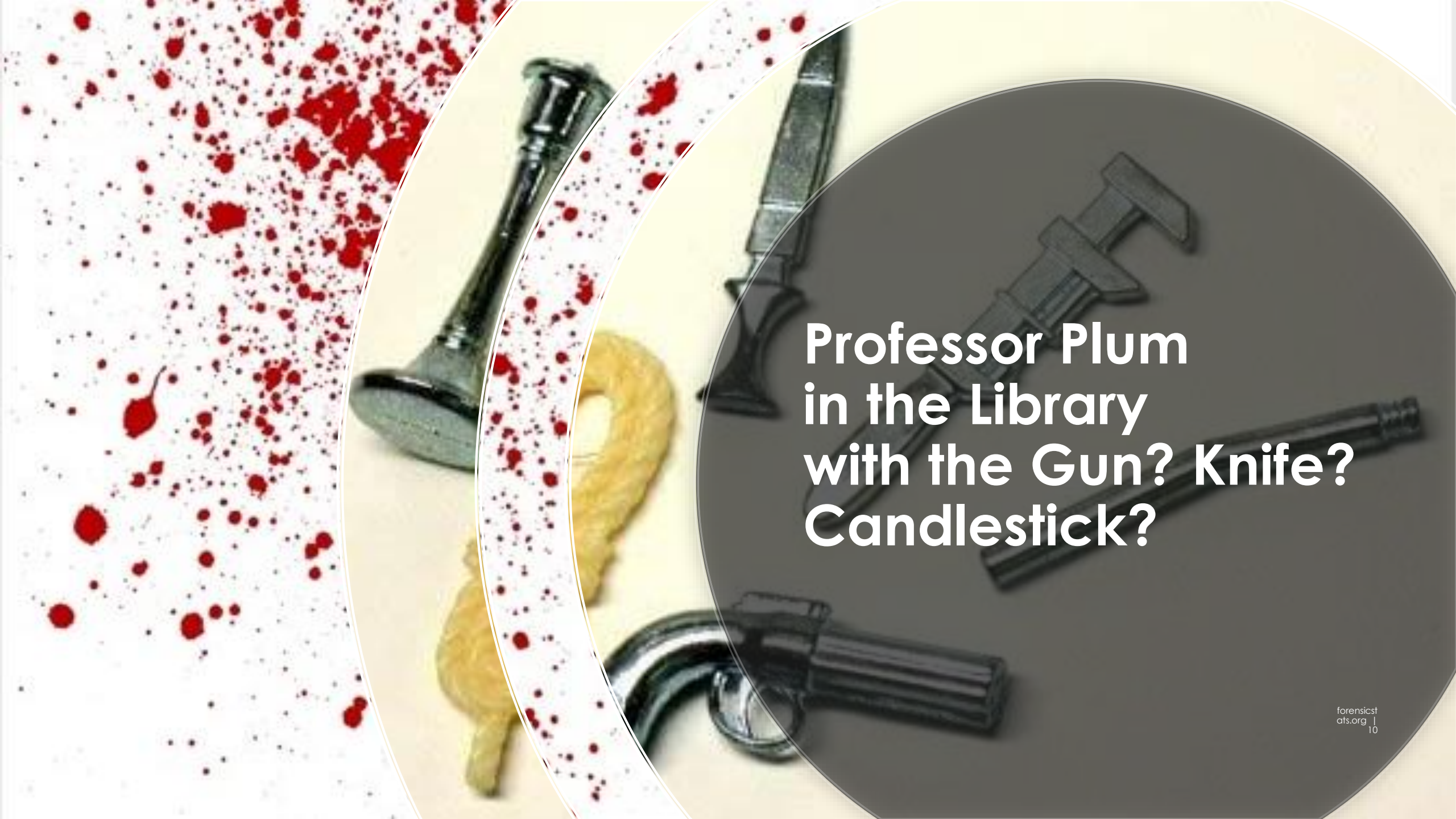R. Austin Hicklin[a,*], Kevin R. Winer[b], Paul E. Kish[c], Connie L. Parks[a], William Chapman[a], Kensley Dunagan[a], Nicole Richetelli[a], Eric G. Epstein[a], Madeline A. Ausdemore[a], Thomas A. Busey[d]

[a] Noblis, Reston, VA, USA
[b] Kansas City Police Department Crime Laboratory, Kansas City, MO, USA
[c] Forensic Consultant, Corning, NY, USA
[d] Indiana University, Bloomington, IN, USA

ARTICLE INFO

ABSTRACT

Although the analysis of bloodstain pattern evidence left at crime scenes relies on the expert opinions of bloodstain pattern analysts, the accuracy and reproducibility of these conclusions have never been rigorously evaluated at a large scale. We investigated conclusions made by 75 practicing bloodstain pattern analysts on 192 bloodstain patterns selected to be broadly representative of operational casework, resulting in 33,005 responses to prompts and 1760 short text responses. Our results show that conclusions were often erroneous and often contradicted other analysts. On samples with known causes, 11.2% of responses were erroneous. The results show limited reproducibility of conclusions: 7.8% of responses contradicted other analysts. The disagreements with respect to the meaning and usage of BPA terminology and classifications suggest a need for improved standards. Both semantic differences and contradictory interpretations contributed to errors and disagreements, which could have serious implications if they occurred in casework.

The process of scientific maturity – Evidence-based Medicine

Levels of Evidence

Randomized Controlled Trials (RCTs) "the gold standard"

PCAST

Traditional Forensic Science

Level I: Evidence from one or more RCTs

Level II-1: Evidence from controlled trials without randomization

Level II-2: Evidence from cohort or case-control analytic studies

Level II-3: Evidence from multiple time series (observational studies)

Level III: Opinions of respected authorities, based on clinical experience, descriptive studies, or reports of expert committees (ideally using formal consensus methods)

Level IV: "Evidence" based on personal anecdote ("In my experience…")

Haughom, J. "Reasons the practice of evidence-Based medicine is a hot topic." (5).

I'm practicing "songwriting" not acoustics.

Metrology?

I'm practicing "soaring" not aerodynamics

Metrology???

The process of scientific maturity – Forensic Science

United States Department of Justice Statement on the PCAST Report: *Forensic Science in Criminal Courts: Ensuring Scientific Validity of Feature-Comparison Methods*

In September 2016, the President's Council of Advisors on Science and Technology ("PCAST") released its report, *Forensic Science in Criminal Courts: Ensuring Scientific Validity of Feature-Comparison Methods.*[1] The stated purpose of the Report was to determine what additional scientific steps could be taken after publication of the 2009 National Research Council Report[2] to ensure the validity of forensic evidence used in the legal system.[3] PCAST identified what it saw as two important gaps: 1) the need for clarity about scientific standards for the validity and reliability of forensic methods; and 2) the need to evaluate specific methods to determine whether they had been scientifically established as valid and reliable.[4] The Report "aimed to close these gaps" for a number of what it described as "feature comparison methods."[5] These are methods for comparing DNA samples, latent fingerprints, firearm marks, footwear patterns, hair, and bitemarks.[6]

Unfortunately, the PCAST Report contained several fundamentally incorrect claims. Among these are: 1) that traditional forensic pattern comparison disciplines, as currently practiced, are part of the scientific field of metrology; 2) that the validation of pattern comparison methods can *only* be accomplished by strict adherence to a non-severable set of experimental design criteria; and 3) that error rates for forensic pattern comparison methods can *only* be established through "appropriately designed" black box studies.

The purpose of this statement is to address these claims and to explain why each is incorrect. After the PCAST Report was released, the Department of Justice ("Department") announced that it would not follow PCAST's recommendations.[7] The Report was criticized by a number of commentators and organizations outside of the Department for its analysis, conclusions, factual inaccuracies, and other mistakes.[8] Formally addressing PCAST's incorrect claims has become

[1] PRESIDENT'S COUNCIL OF ADVISORS ON SCI. & TECH., EXEC. OFFICE OF THE PRESIDENT, FORENSIC SCI. IN CRIM. COURTS: ENSURING SCI. VALIDITY OF FEATURE COMPARISON METHODS (2016), https://obamawhitehouse.archives.gov/sitesdefault/files/microsites/ostp/PCAST/pcast_forensic_science_report_final [https://perma.cc/VJB4-5JVQ] [hereinafter PCAST REPORT].
[2] NAT'L RES. COUNCIL, NAT'L ACAD'S., STRENGTHENING FORENSIC SCI. IN THE UNITED STATES: A PATH FORWARD 122 (Nat'l Acad. Press 2009).
[3] PCAST REPORT, *supra* note 1, at 1.
[4] *Id.*
[5] *Id.* In this statement, we use the term "pattern comparison," rather than PCAST's chosen term, "feature comparison" to describe the general nature of the methods discussed.
[6] *Id.* Department of Justice laboratories do not practice what PCAST described as "bitemark analysis."
[7] Gary Fields, *White House Advisory Council Is Critical of Forensics Used in Criminal Trials*, WALL ST. J. (Sept. 20, 2016, 4:25 PM), https://www.wsj.com/articles/whitehouse-advisory-council-releases-report-critical-of-forensics-used-in-criminal-trials-a1474394743 [https://perma.cc/N9KM-NHJL].
[8] *See, i.e.,* I.W. Evett et al., *Finding a Way Forward for Forensic Science in the US—A Commentary on the PCAST Report*, 278 FORENSIC SCI. INT'L 16, 22–23 (2017); Letter from Michael A. Ramos, President, Nat'l Dist. Attorneys Ass'n, to President Barack Obama (Nov. 16, 2016), http://tinyurl.com/hczkt3k.; Ass'n of Firearms and Toolmark Examiners (AFTE) Response to PCAST Report on Forensic Sci. (October 31, 2016), https://afte.org/uploads/documents/AFTE-PCAST-Response.pdf; Org. of Sci. Area Committees (OSAC) Firearms and Toolmarks Subcommittee Response to the President's Council of Advisors on Sci. and Tech. (PCAST) Call for Additional References Regarding its Rep. "Forensic Sci. in Crim. Courts: Ensuring Sci. Validity of Feature-

1

The US Department of Justice stumbles on visual perception

Thomas D. Albright[a,1]

A large and highly valuable category of forensic evidence consists of patterned impressions created during the perpetration of a crime. These crime scene artifacts, such as fingerprints or tire tracks, offer visual sensory information that is assessed by trained human observers and compared to sensory experiences elicited by model patterns that would have been produced under a hypothesized set of conditions. By means of this "forensic feature comparison," the observer makes a judgment about whether the evidence and the model are sufficiently similar to support common origin. In light of documented failures of this approach, significant concerns have been raised about its scientific validity. In response to these concerns, the US Department of Justice has made assertions about how forensic examiners perform feature comparison tasks that are not consistent with modern scientific understanding of the processes of sensation and perception. Clarification of these processes highlights new ways of thinking about and improving the accuracy of forensic feature comparison and underscores the vital role of science in achieving justice.

forensic science | sensory measurement | feature comparison

Forensic science is the bread and butter of criminal investigation and prosecution. On the surface of things, it is an incredibly compelling discipline. Artifacts of human activity left without intent or awareness suggest specific action scenarios, implicate specific actors, and sometimes support inferences regarding an actor's motivation or intent. Indeed, much of the genuine public fascination with—and trust in—forensic science stems from the sense of eavesdropping, from the feeling that we might learn some raw truth that is infinitely more candid simply because the actor was unaware of being watched. All of these fuel righteous indignation against those who would cause criminal offense and gives us the satisfying impression that we have a leg up on the bad guys.

Despite this fanciful optimism and longstanding public support, it has become increasingly clear that forensic practices that rely on human judgment often implicate the wrong people. This form of error frequently has tragic personal and societal consequences, including wrongful conviction and imprisonment. Indeed, thousands of innocent person-years have been spent behind bars for this reason, the majority of these quashed lives being men of color (1).

Many of the problems with this discipline were considered in a landmark 2009 report from the National Academy of Sciences (NAS) (2). This congressionally mandated study identified numerous weaknesses associated with validation, training, and reporting procedures in forensic practice and included detailed recommendations for science-based reform. These recommendations led, most notably, to creation of the short-lived National Commission on Forensic Science, and the National Institute of Standards and Technologies operation known as the Organization of Scientific Area Committees for Forensic Science, and to a variety of grass-roots efforts to improve and standardize forensic practice. In 2015, President Obama asked the President's Council of Advisors on Science and Technology (PCAST) to further evaluate needs within the forensic science community, the product of which was a 2016 report focusing on a specific subset of forensic practices known as "feature comparison" methods (3).

Forensic Feature Comparison

Feature comparisons are among the oldest and most commonly employed of forensic methods and are familiar to most by their use for evaluation of visually patterned

[a]The Salk Institute for Biological Studies, La Jolla, CA 92037
Author contributions: T.D.A. wrote the paper.
The author declares no competing interest
This article is a PNAS Direct Submission.

[1]Email: tom@salk.edu.

Applying Algorithms to Forensic Science

Jeff Salyards, PhD, MFS
Forensic Research Scientist IV

Center for Statistics and Applications in
Forensic Evidence (CSAFE)
Iowa State University