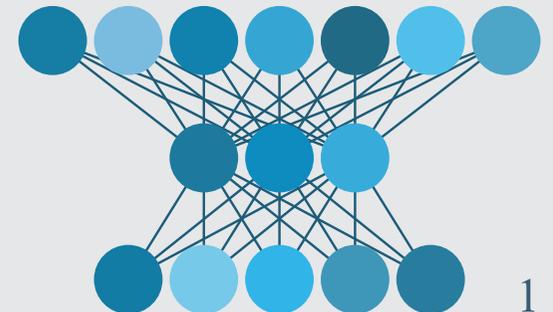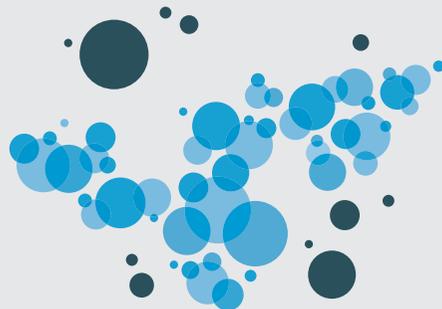# Advancing a paradigm shift in evaluation of forensic evidence

Geoffrey Stewart Morrison

Forensic Data Science Laboratory
Aston Institute for Forensic Linguistics

Aston University
BIRMINGHAM UK

1

Acknowledgment

Disclaimers

- All opinions expressed are those of the presenter and, unless explicitly stated otherwise, should not be construed as representing the policies or positions of any organizations with which the presenter is associated.

- This presentation is based on part of a paper that is soon to be submitted. Claims made in this presentation are supported by additional references which are cited in the paper, but which have been omitted from these slides.

Contents

- Status quo

- Quo vadis?

- Impedimenta

- Via progredi

Status quo

Status quo

- Curran (2013):

    Is forensic science the last bastion of resistance against statistics?

Status quo

- UK House of Lords Science and Technology Committee (2019):

  In regard to pattern comparison methods, ... "the comparison of fingerprints, toolmarks, footwear, tire marks and ballistics" [are] "spot-the-difference" techniques in which "there is little, if any, robust science involved in the analytical or comparative processes used and as a consequence there have been questions raised around the reproducibility, repeatability, accuracy and error rates of such analysis." (§155)

Status quo

- The process of **evaluation of strength of forensic evidence** consists of:

    - **analysis**, i.e., extraction of information from items of interest (the evidence); and

    - **interpretation**, i.e., drawing inferences with respect to the meaning of the information extracted by the analysis.

Status quo

- Items of interest may be, for example:

    - a **fingermark** of questioned source recovered from a crime scene and a fingerprint collected from a known individual;

    - a **recording of a speaker** of questioned identity on an intercepted telephone call and a recording of a police interview with a speaker of known identity;

    - a **fired cartridge case** recovered from a crime scene and cartridge cases fired in a forensic laboratory from a gun found in the possession of a suspected shooter.

Status quo

- Forensic practitioners conduct evaluations in order to **assist legal-decision makers to make decisions with respect to questions** of legal concern such as:

  - Do the fingermark and fingerprint originate from the same finger?

  - Is the speaker of questioned identity on the intercepted recording the same as the speaker of known identity?

  - Was the cartridge case recovered from the crime scene fired from the suspect's gun?

Status quo

- Currently, across the majority of branches of forensic science, widespread practice is that:

  - **analysis** is conducted using **human perception**, and

  - **interpretation** is conducted using **subjective judgement**.

- Even in branches of forensic science in which analysis is conducted using instrumental measurement, interpretation is commonly based on subjective judgement, e.g., by eyeballing graphical representations of the measured values.

Status quo

- Human-perception-based analysis methods and subjective-judgement-based interpretation methods are:

    - **non-transparent** and

    - **susceptible to cognitive bias**.

Status quo

- Across the majority of branches of forensic science, even branches of forensic science in which interpretation is conducted using statistical models:

    - interpretation of evidence is often **logically flawed**, and

    - forensic-evaluation systems (the end-to-end combination of analysis and interpretation methods) are often **not empirically validated** or not adequately empirically validated.

Quo vadis?

Quo vadis?

- Saks & Koehler (2005):

  we envision a **paradigm shift** in the traditional forensic identification sciences in which **untested assumptions and semi-informed guesswork** are **replaced by a sound scientific foundation and justifiable protocols**. Although obstacles exist both inside and outside forensic science, the time is ripe for the traditional forensic sciences to **replace antiquated assumptions of uniqueness and perfection with a more defensible empirical and probabilistic foundation**. (p. 895)

Quo vadis?

- US President's Council of Advisors on Science and Technology (PCAST, 2016):

  neither experience, nor judgment, nor good professional practice ... can substitute for **actual evidence of foundational validity and reliability**. The frequency with which a particular pattern or set of features will be observed in different samples, which is an essential element in drawing conclusions, is not a matter of "judgment." It is an empirical matter for which **only empirical evidence is relevant**. (p. 6)

Quo vadis?

- US President's Council of Advisors on Science and Technology (PCAST, 2016):

  Objective methods are, in general, preferable to subjective methods. Analyses that depend on human judgment (rather than a quantitative measure ...) are obviously more susceptible to human error, bias, and performance variability across examiners. In contrast, objective, quantified methods tend to yield greater accuracy, repeatability and reliability, including reducing variation in results among examiners. **Subjective methods can evolve into or be replaced by objective methods**. (p. 47)

Quo vadis?

- A **paradigm shift** in evaluation of forensic evidence in which methods based on human perception and subjective judgement are replaced by methods based on **relevant data**, **quantitative measurements**, and **statistical models**; methods that:

  - are **transparent and reproducible**;

  - are **intrinsically resistant to cognitive bias**;

  - use the **logically correct framework for interpretation of evidence** (the **likelihood-ratio framework**); and

  - are **empirically validated under casework conditions**.

Quo vadis?

- **Transparency and reproducibility:**

  - Methods dependent on human perception and subjective judgement are intrinsically non-transparent and therefore not reproducible by others.

  - **Human introspection is often mistaken**, hence a forensic practitioner's explanation of how they reached their conclusion may not reflect how they actually reached that conclusion.

Quo vadis?

- **Transparency and reproducibility:**

    - In contrast, procedures based on data, quantitative measurement, and statistical models are transparent and reproducible.

    - **Measurement (feature-extraction) and statistical-modelling methods can be described in detail, and data and software tools can potentially be shared with others.**

Quo vadis?

- **Cognitive bias:**

  - Cognitive bias is subconscious bias, it cannot be controlled by strength of will.

  - **Forensic practitioners are susceptible to cognitive bias when making perceptual observations**: their belief in the probability that a hypothesis is true can affect their analysis of the evidence and therefore the information that feeds into their interpretation.

Quo vadis?

- **Cognitive bias:**

  - Cognitive bias is subconscious bias, it cannot be controlled by strength of will.

  - **Forensic practitioners are susceptible to cognitive bias when they are making subjective judgements** and are exposed to information that could influence their belief in the probability that a hypothesis is true but that would not logically affect the probability of obtaining the evidence conditional on whether the hypothesis were true.

Quo vadis?

- **Cognitive bias:**

  - Cognitive bias is subconscious bias, it cannot be controlled by strength of will.

  - **Systems in which the strength-of-evidence conclusion is directly the result of subjective judgement are particularly susceptible to cognitive bias.**

Quo vadis?

- **Cognitive bias:**

  - Systems based on quantitative measurements and statistical models require subjective judgements on decisions such as:

    - whether the **data used for training the system** are sufficiently representative of the relevant population for the case and sufficiently reflective of the conditions of the items of interest in the case so that the output of the system will be a meaningful answer to the question posed in the case.

Quo vadis?

- **Cognitive bias:**

    - Systems based on quantitative measurements and statistical models require subjective judgements on decisions such as:

        - whether the **data used for validating the system** are sufficiently representative of the relevant population for the case and sufficiently reflective of the conditions of the items of interest in the case so that the results of the validation will provide a meaningful indication of the performance of the systems under the conditions of the case.

Quo vadis?

- **Cognitive bias:**

    - Systems based on quantitative measurements and statistical models require subjective judgements.

    - These judgements, however, are **made at the beginning of the process** before the practitioner has analyzed the items of interest, hence the practitioner cannot know what effect these decisions will have on the strength-of-evidence conclusion.

    - **The remainder of the evaluation process is automated**, hence not susceptible to cognitive bias.

Quo vadis?

- **Likelihood-ratio framework:**

  - In current practice, interpretation of evidence is often logically flawed, e.g.:

    - it is based on the **uniqueness** or **individualization fallacy**,

    - and conclusions are often expressed **categorically**, e.g., "identification", "inconclusive", "exclusion",

    - or using some form of **uncalibrated verbal posterior-probability scale**, e.g., "identification", "probable identification", "inconclusive", "probable exclusion", "exclusion".

Quo vadis?

- **Likelihood-ratio framework:**

  - In contrast, the likelihood-ratio framework is **advocated as the logically correct framework** for evaluation of evidence **by the vast majority of experts in forensic inference and statistics**, including:

    - Aitken et al. (2011) with 31 authors/supporters

    - Morrison et al. (2017) with 19 authors/supporters

    - Morrison, Enzinger, et al. (2021) with 20 authors/supporters

Quo vadis?

- **Likelihood-ratio framework:**

  - In contrast, the likelihood-ratio framework is **advocated as the logically correct framework** for evaluation of evidence **by key organizations**, including:

    - Association of Forensic Science Providers of the United Kingdom and of the Republic of Ireland (AFSP)

    - Royal Statistical Society (RSS)

    - European Network of Forensic Science Institutes (ENFSI)

    - National Institute of Forensic Science of the Australia New Zealand Policing Advisory Agency (NIFS)

    - American Statistical Association (ASA)

    - Forensic Science Regulator for England & Wales (FSR)

Quo vadis?

- **Likelihood-ratio framework:**

    - The likelihood-ratio framework requires assessment of the **probability of obtaining the evidence if one hypothesis were true versus the probability of obtaining the evidence if an alternative hypothesis were true**.

    - The **two hypotheses** must be **mutually exclusive**.

    - One hypothesis should represent the **position of the prosecution** in the case, and the other the **position of the defence**.

Quo vadis?

- **Likelihood-ratio framework:**

  - e.g., the fingermark of questioned origin was deposited by a finger of a particular known individual, versus the fingermark of questioned origin was deposited by a finger of some other individual selected at random from the relevant population,

  - for which the **numerator** of the likelihood ratio **quantifies the** *similarity* between the mark and the print, and the **denominator quantifies the** *typicality* of the mark **with respect to the relevant population**.

Quo vadis?

- **Empirical validation:**

  - **Empirical validation under conditions reflecting those of the case** to which a forensic-evaluation system is to be applied is the **only way to know how well that system performs under the conditions of the case**.

  - Over the last 20 years, protocols for validating systems that output likelihood ratios have been developed, including metrics and graphics appropriate for representing the results of such validations.

Quo vadis?

- **Empirical validation:**

  - The need for validation under casework conditions has been emphasized by FSR (2020b), and by PCAST (2016):

    Without appropriate estimates of accuracy, an examiner's statement that two samples are similar—or even indistinguishable—is scientifically meaningless: it has no probative value, and considerable potential for prejudicial impact. Nothing—not training, personal experience nor professional practices—can substitute for adequate empirical demonstration of accuracy. (p. 46)

Quo vadis?

- **Empirical validation:**

  - Despite this, practitioners in multiple branches of forensic science:

    - often claim that training and experience provide sufficient warrant for their conclusions

    - deny or obfuscate about the need for validation

    - propose lax validation protocols that do not require demonstration of performance under casework conditions

Quo vadis?

- **A Kuhnian paradigm shift:**

  - The most famous article heralding a paradigm shift in evaluation of forensic evidence is Saks & Koehler (2005).

  - I believe that Saks & Koehler and I describe the same paradigm shift.

  - Saks & Koehler stated that they intended "paradigm shift" as a metaphor, however, ...

Quo vadis?

- **A Kuhnian paradigm shift:**

  - I view the paradigm shift in evaluation of forensic evidence as a true Kuhnian paradigm shift (Kuhn, 1962) in the sense that:

  - It **requires rejection of existing methods** and the **ways of thinking** that underpin them,

  - and **rejection of the idea that progress can be made by incremental improvements** to existing methods.

  - It **requires the wholesale adoption of an entire constellation of new methods and new ways of thinking**.
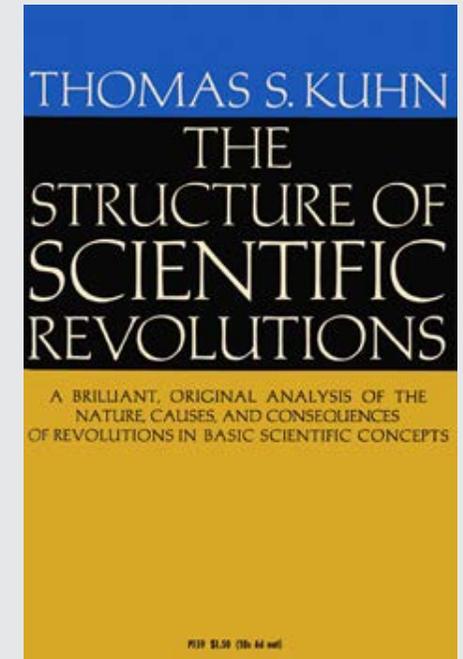
Impedimenta

Impedimenta

- The paradigm shift in evaluation of forensic evidence is ongoing, but progress is slow or stalling for the following reasons:

- **The new paradigm has only been adopted in a few branches of forensic sciences, and only by a minority of researchers and practitioners.**

  - DNA

  - forensic voice comparison

Impedimenta

- **Only some elements of the new paradigm have been adopted as part of incremental change.**

    - Kuhn (1962) p. 149:

        Just because it is a transition between incommensurables, the transition between competing paradigms cannot be made a step at a time, ... Like the gestalt switch, it must occur all at once (though not necessarily in an instant) or not at all.

Impedimenta

- **There is misunderstanding of the new paradigm and resistance to its adoption.**

  - The cultures of some branches of forensic science seem to be especially resistant to the adoption of statistical-model-based methods and of validation.

  - There are many examples of legal rulings in which judges have misunderstood the meaning of a likelihood ratio or have not understood empirical validation and its importance.

Impedimenta

- **Research is often not informed by practice and has no impact on practice.**

    - Margot (2011) p. 801:

        Research in forensic science is sorely needed, but it should address primarily forensic science questions—not questions relating to the application of chemistry, biology, statistics, or psychology.

Impedimenta

- **It is difficult to obtain funding for evidential-forensic-science research.**

    - Roux et al. (2021) p. 679:

        technology-oriented development ... often overrul[es] the importance of appropriate scientific reasoning to solve actual problems

    - House of Lords (2019) §187 recommended that UKRI urgently and substantially increase the amount of dedicated funding allocated to forensic science

Impedimenta

- **There are genuine practical impediments to implementing the new paradigm.**

    - Even if practitioners want to adopt the new paradigm, they will be unable to do so unless they are provided with:

        - quantitative-measurement and statistical-modelling tools and case-relevant data necessary for calculating likelihood ratios and for validating system performance

        - training on understanding the new paradigm and on how to implement it for casework

Via progredi

Via progredi

- Kuhn (1962):

  The transfer of allegiance from paradigm to paradigm is a conversion experience that cannot be forced. ... a generation is sometimes required to effect the change ... Conversions will occur a few at a time until, after the last holdouts have died, the whole profession will again be practicing under a single, but now a different, paradigm. (pp. 150–151)

Via progredi

- My **strategy** is to work with researchers and practitioners <u>who want to adopt the new paradigm</u>, to work with them on addressing practical impediments to applying the new paradigm in casework:

    - to provide researchers, practitioners, and lawyers with training leading to understanding of the new paradigm

    - to collaborate with researchers and practitioners on building relevant databases and on developing and validating statistical models applicable in their particular branches of forensic science

Via progredi

- My **strategy** is to work with researchers and practitioners <u>who want to adopt the new paradigm</u>, to work with them on addressing practical impediments to applying the new paradigm in casework:

    - to conduct research on how to present likelihood ratios and validation results so as to maximize understanding by laypeople, and thereby provide guidance to forensic practitioners on how to communicate forensic-evaluation results to legal-decision makers

Via progredi

- My **strategy** is to work with researchers and practitioners <u>who want to adopt the new paradigm</u>, to work with them on addressing practical impediments to applying the new paradigm in casework:

    - build on knowledge gained from the experience of advancing the paradigm shift in forensic voice comparison

    - including transferring and adapting statistical modelling techniques, and calibration and validation procedures, used in forensic voice comparison

Via progredi

Ongoing programmes of work:

- **Calibration and validation** (past 14 years)

  - Applicable across multiple branches of forensic science

- **Forensic voice comparison** (past 14 years)

  - Phil Weber, Aston Forensic Data Science Laboratory

  - Multiple external collaborators and partner organizations

  - State-of-the-art automatic-speaker-recognition technology

- **Fired-cartridge-case comparison** (past 2 years)

  - Nabanita Basu, Aston Forensic Data Science Laboratory

  - Rachel S Bolton-King, Staffordshire University

Via progredi

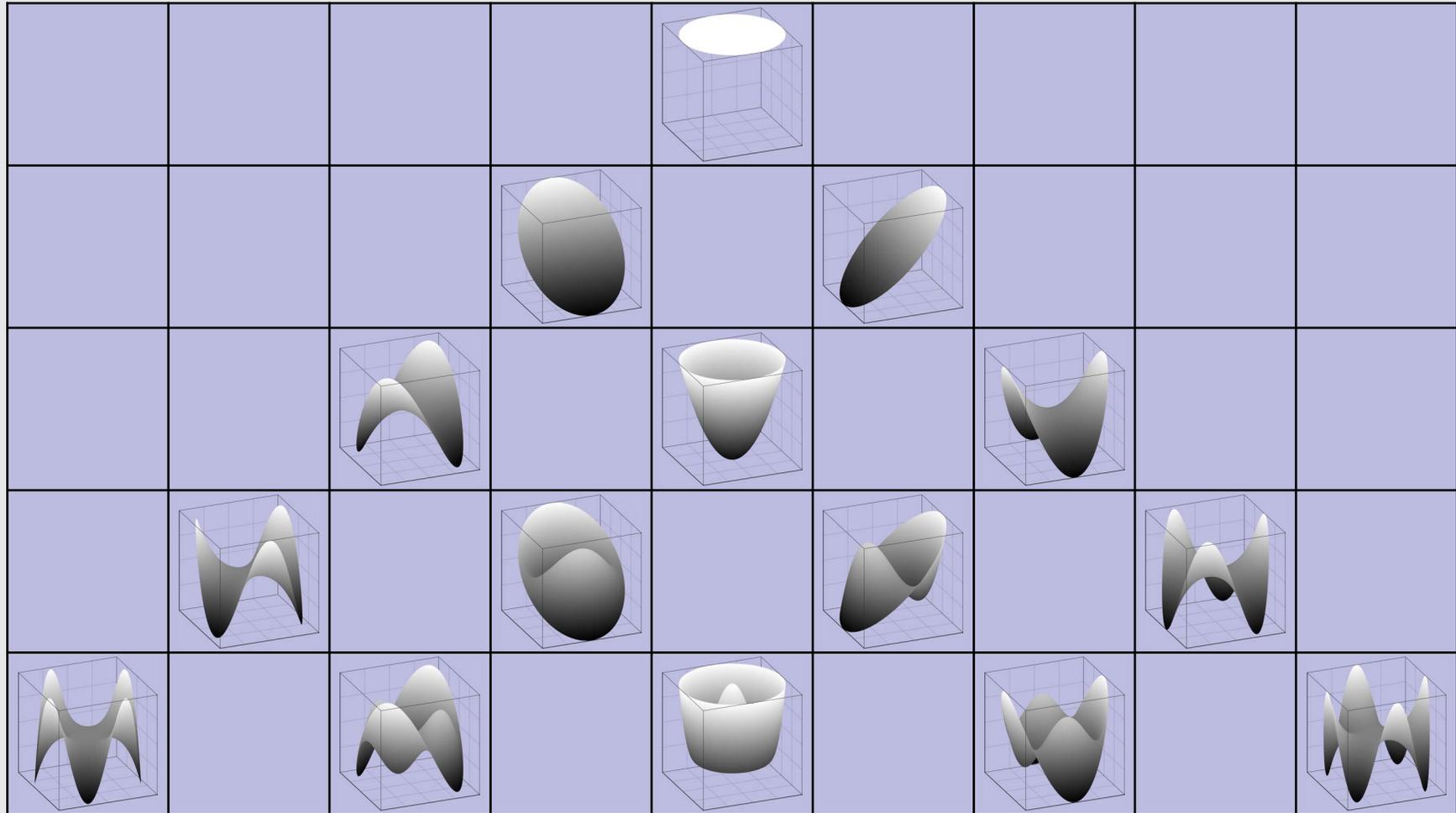Ongoing programmes of work:

- **Forensic Anthropology** (within past year)

    - Patrick S Randolph-Quinney, Northumbria University

- **Cell-site analysis** (within past year)

    - Matt Tart, CCL Forensics

- **Communication of forensic science** (within past year)

    - Kristy A Martire & Agnes Bali, University of New South Wales

Via progredi

Potential future programmes of work:

- **Authorship analysis** (beginning)

    - Krzysztof Kredens & Piotr Pezik, Aston Institute for Forensic Linguistics

    - Shunichi Ishihara, Australian National University

- **Fingerprints**

    - grant application in preparation

- **Gait analysis**

    - on hold

- **Other**

    - open to new collaborations with researchers and practitioners

Thank You

http://geoff-morrison.net/