

Mesa Redonda

Bayesian Inference in Speaker Recognition

Geoffrey Stewart Morrison

$$\frac{p(E|H_p)}{p(E|H_d)}$$

- **CONCERNS**

- **Logically correct framework for evaluation of forensic evidence** (ENFSI Guideline for Evaluative Reporting 2015)
- **But what is the warrant for the opinion expressed? Where do the numbers come from?** (Risinger at ICFIS 2011)
- **Demonstrate validity and reliability** (*Daubert* 1993; National Research Council 2009; Forensic Science Regulator Codes of Practice 2014)
- **Transparency** (*R v T* 2010)
- **Reduce potential for cognitive bias** (NIST/NIJ Human Factors in Latent Fingerprint Analysis 2012)

- **PARADIGM**

- **Use of likelihood ratio framework**

- Logically correct framework for evaluation of evidence.
- Specific prosecution and defence hypotheses adopted by forensic scientist must be explained to judge at admissibility hearing / trier of fact at trial.
- Is the question appropriate?
- Question must be understood in order to understand answer.

- **PARADIGM**

- **Calculation of numeric likelihood ratios using relevant data, quantitative measurements, and statistical models**
 - Sample from the relevant population specified in the defence hypothesis. Sufficiently representative?
 - Data reflective of conditions of suspect and offender samples. Sufficiently reflective?
 - Report output of statistical model, keep subjective elements far from the conclusion. Do not report conclusions which are primarily or directly based on subjective judgement.

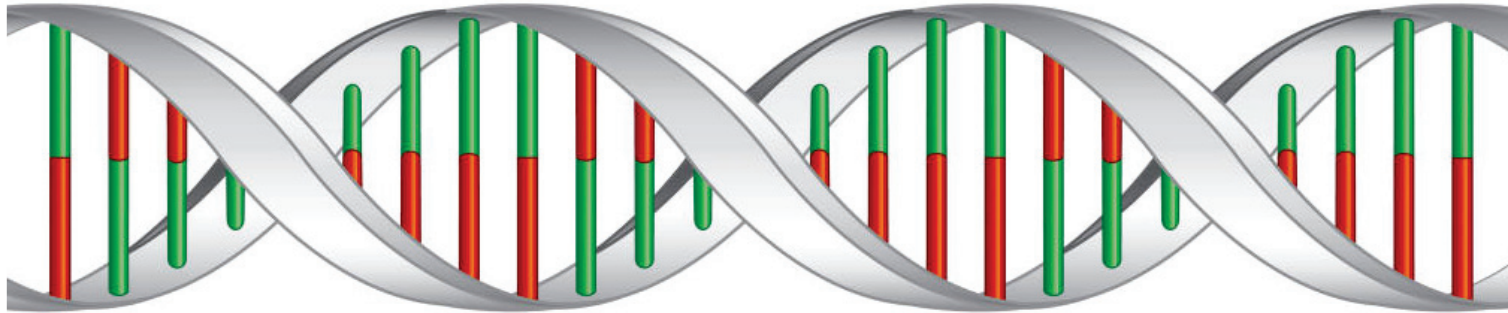
- **PARADIGM**

- **Empirical testing of validity and reliability under conditions reflecting those of the case under investigation**

- Performance under different conditions may be very different.
- Sample from the relevant population specified in the defence hypothesis. Sufficiently representative?
- Data reflective of conditions of suspect and offender samples. Sufficiently reflective?
- Test the system actually employed, including human operator.

Likelihood ratios

- Adopted as standard for evaluation of DNA evidence in mid 1990's



Likelihood ratios



- Association of Forensic Science Providers (2009)
 - *Standards for the formulation of evaluative forensic science expert opinion*
- 31 signatories [from Aitken to Zadora] (2011)
 - *Expressing evaluative opinions: A position statement*
- European Network of Forensic Science Institutes (2015)
 - *Guideline for evaluative reporting in forensic science*



Likelihood ratios

- Morrison (2010)

- Morrison, G. S. (2010). Forensic voice comparison. In I. Freckelton, & H. Selby (Eds.), *Expert Evidence* (Ch. 99). Sydney, Australia: Thomson Reuters.

- <http://expert-evidence.forensic-voice-comparison.net/>

- Morrison (2012)

- Morrison, G. S. (2012). The likelihood-ratio framework and forensic evidence in court: A response to R v T. *International Journal of Evidence and Proof*, 16, 1–29. doi:10.1350/ijep.2012.16.1.390

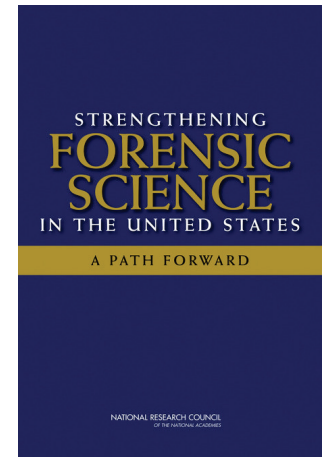
Calculate and report numeric likelihood ratios

- Morrison & Stoel (2014)

- Morrison, G. S., & Stoel, R. D. (2014). Forensic strength of evidence statements should preferably be likelihood ratios calculated using relevant data, quantitative measurements, and statistical models – a response to Lennard (2013) Fingerprint identification: How far have we come? *Australian Journal of Forensic Sciences*, 46, 282–292. doi:10.1080/00450618.2013.833648

Validity and Reliability

- The National Research Council report to Congress on *Strengthening Forensic Science in the United States* (2009) urged that procedures be adopted which include:
 - “quantifiable measures of the reliability and accuracy of forensic analyses” (p. 23)
 - “the reporting of a measurement with an interval that has a high probability of containing the true value” (p. 121)
 - “the conducting of validation studies of the performance of a forensic procedure” (p. 121)



Validity and Reliability

- The Forensic Science Regulator of England & Wales' *Codes of Practice and Conduct* (2014) require:
- “all technical methods and procedures used by a provider shall be validated.” (§20.1.1)
- “Even where a method is considered standard and is in widespread use, validation will still need to be demonstrated.” (§20.1.3)
- “validation shall be carried out using simulated casework material ... and ... where appropriate, with actual casework material” (§20.7.3)
- “demonstrate that they can provide consistent, reproducible, valid and reliable results” (§20.9.1)

Validity and Reliability

- Morrison (2014)

- Morrison, G. S. (2014). Distinguishing between forensic science and forensic pseudoscience: Testing of validity and reliability, and approaches to forensic voice comparison. *Science & Justice*, 54, 245–256. doi:10.1016/j.scijus.2013.07.004

- Morrison (2011)

- Morrison, G. S. (2011). Measuring the validity and reliability of forensic likelihood-ratio systems. *Science & Justice*, 51, 91–98. doi:10.1016/j.scijus.2011.03.002

Demonstrations of implementing paradigm

- Enzinger et al. (2015)

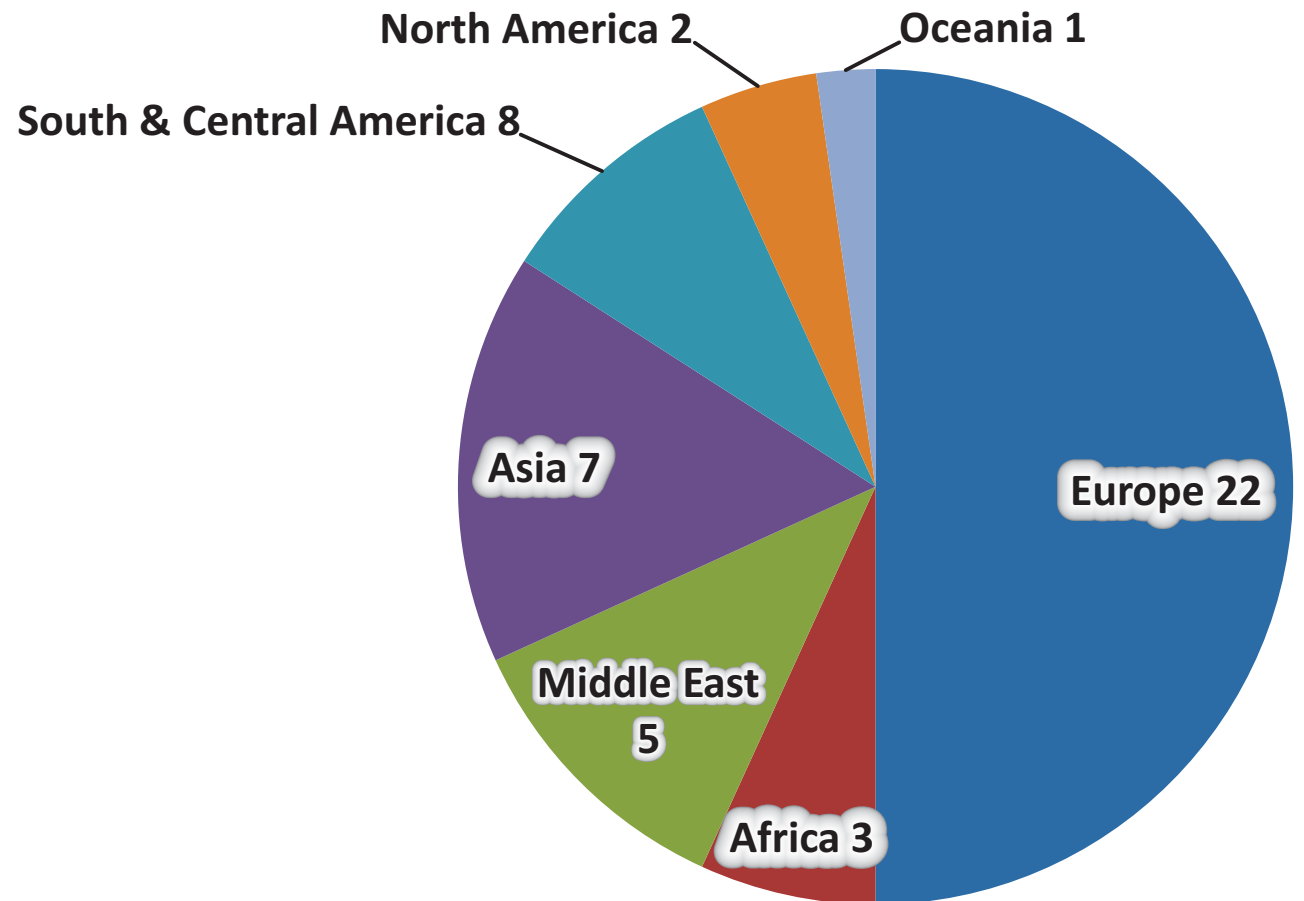
- Enzinger, E., Morrison, G. S., & Ochoa, F. (2015). A demonstration of the application of the new paradigm for the evaluation of forensic evidence under conditions reflecting those of a real forensic-voice-comparison case. *Science & Justice*. doi:10.1016/j.scijus.2015.06.005

- Enzinger & Morrison (2015)

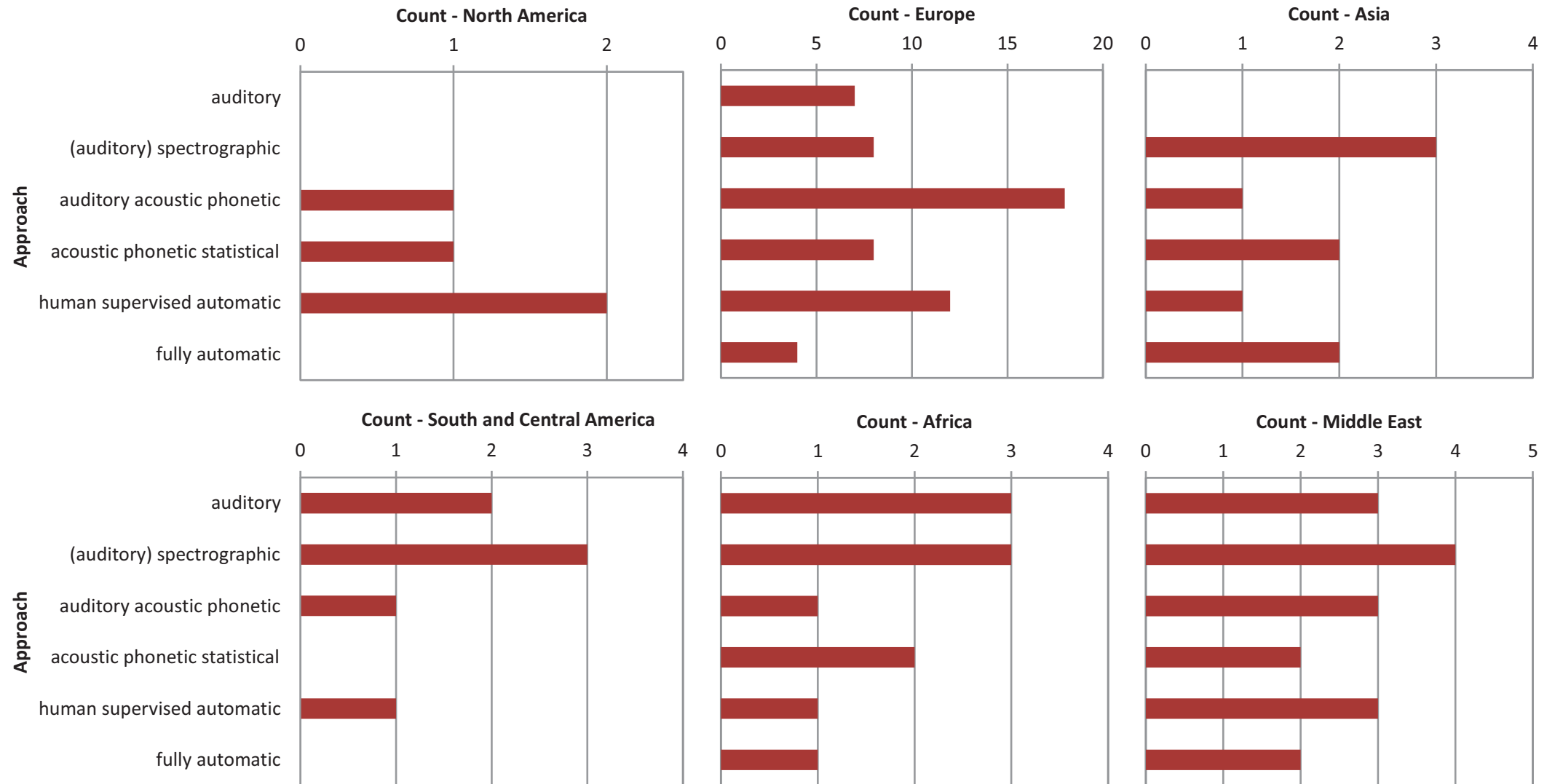
- Enzinger, E., & Morrison, G. S. (2015). Mismatched distances from speakers to telephone in a forensic-voice-comparison case. *Speech Communication*, 70, 28–41. doi:10.1016/j.specom.2015.03.001

INTERPOL Survey

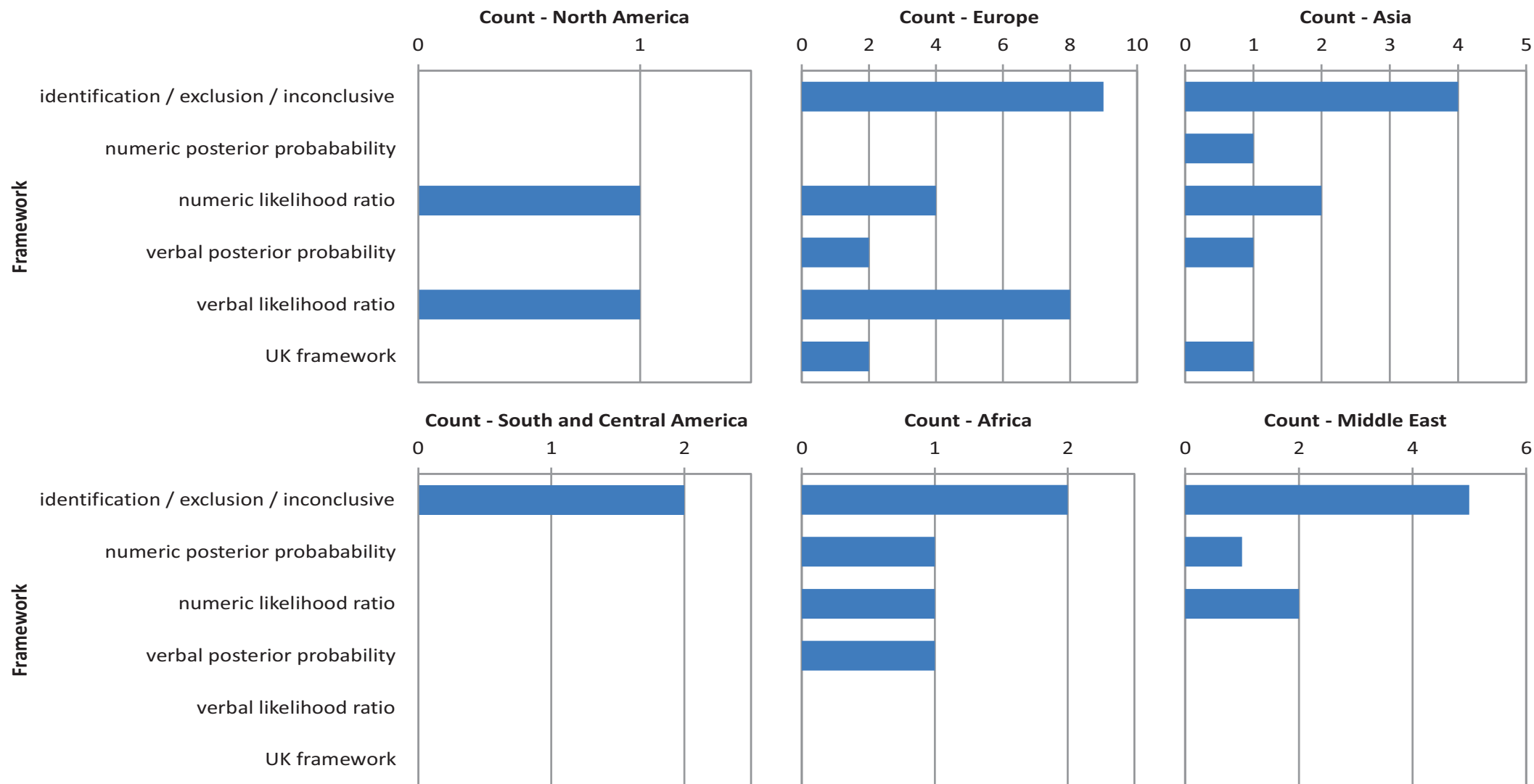
- Number of responses:
 - 91 from 69 countries
 - 44 had speaker identification capabilities in house or via external laboratories



INTERPOL Survey



INTERPOL Survey



Thank You

<http://geoff-morrison.net/>

<http://forensic-evaluation.net/>