

## **A response to the discussion of score-based models in Aitken (2018) “Bayesian hierarchical random effects models in forensic science”**

Dear Editor:

In the process of writing a summary article, one must omit details. Unfortunately, this process can sometimes inadvertently lead to information being presented in a way that is potentially misleading. I write to clarify part of the discussion in Aitken (2018) in the hope of reducing the potential for misunderstandings. Aitken (2018) §3.3 discusses score-based models. It states that “Score-based approaches have been used for handwriting and speech recognition.” The latter term should not be to “speech recognition” (What was said?), but “speaker recognition” (Who was speaking?). It goes on to provide examples of scores that are similarity-only scores – they only measure the degree of similarity (or its inverse, the degree of difference) between each member of a pair of objects – the examples given are Euclidian distance, Manhattan distance, and Pearson correlation distance. Section 3.4 ends by listing shortcomings of the score-based approaches, including: “The value of the likelihood ratio is based on the similarity of pairwise scores rather than the similarity and rarity of features.” In fact, in the field of forensic speaker recognition (aka forensic voice comparison), the scores used are not similarity-only scores, but take account of both similarity and typicality (the latter being the inverse of “rarity”). Morrison & Enzinger (2018) reviewed and illustrated some of the problems of using similarity-only scores, and recommended that scores be based on both similarity and typicality, as is the norm in forensic speaker recognition.

Sincerely

*Geoffrey Stewart Morrison*

Associate Professor of Forensic Speech Science  
Centre for Forensic Linguistics, Aston University, Birmingham, United Kingdom

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

### **References**

- Aitken C.G.G. (2018). Bayesian hierarchical random effects models in forensic science. *Frontiers in Genetics*, article 126.  
<https://doi.org/10.3389/fgene.2018.00126>
- Morrison G.S., Enzinger E. (2018). Score based procedures for the calculation of forensic likelihood ratios – Scores should take account of both similarity and typicality. *Science & Justice*, 58, 47–58.  
<https://doi.org/10.1016/j.scijus.2017.06.005>