Letter to the Editor:

Reply to Morrison et al. (2016) Refining the relevant population in forensic voice comparison – A response to Hicks et alii (2015) The importance of distinguishing information from evidence/observations when formulating propositions

T. Hicks, A. Biedermann, J.A. de Koeijer, F. Taroni, C. Champod, I.W. Evett

Dear Editor,

In their letter to the Editor, Morrison et al. [1] react to our paper [2] on the importance of distinguishing information from evidence when formulating propositions. We wish to thank the authors for this discussion, as it is essential for readers to understand that our paper [2] is not meant to be prescriptive. The aim of our article is to underline the importance of taking sufficient time to define precisely what is the evidence the forensic scientist is asked to assess and what are the information and the propositions, the latter being assessed by the Court. Once alerted, the expert will take the appropriate choice. Thus, our paper should be seen as a framework for thinking, and not as a list of default propositions.

The formulation of propositions is not only very difficult but it is also at the core of the evaluative process. Through involvement in education [3], we have had many opportunities to discuss this important topic, and that is why we consider it valuable to share our experience with forensic practitioners.

The first point of our paper [2] was to show that in many instances, across forensic science disciplines, observations made on the examined questioned item(s) are mixed with propositions – in particular in the alternative proposition. When doing so, part of the evidence is built into the proposition and hence cannot be formally assessed by the scientist anymore. The reason for this is that scientists assess the evidence and the court assesses propositions. This also holds for forensic speech analysts, as reaffirmed by Morrison et al. in [1]. Thus, it is imperative to detect the problem and clarify it for the judiciary.

Our second point was to emphasize that *if* propositions incorporate aspects that relate to the observations, then the scientist's evaluation of the value of the forensic findings will answer a different question than in a case in which the definition of the propositions is not based on the observations. This is an entirely general point that, again, holds for forensic comparative analyses of speech, too, agreed also by Morrison et al. [1] and illustrated in terms of a practical example.

Specifically, Morrison et al. [1] underline that technical systems and associated evaluative approaches may work better in cases in which propositions are defined based on some observations made on the questioned item(s).

As we will show, this is not problematic, and one can cope with this aspect without changing the propositions (which are to be based on case information and not on forensic observations).

We will use a little notation, as this helps formalisation. Let us denote by 'E' the evidence that forensic scientists or experts in forensic voice comparison need to assess. Our short speech example was as follows: *in a case involving the recording of a phone call, the police officer in charge of the case believes that Mr B has a voice that sounds like the offender's voice heard on*

the recording, because of the notable regional accent. She requests a speech analysis of the recording. We noted that *'If the observations of the police officer (i.e., the offender appeared to have an accent from a given region) are not contested by the defence and can be taken as a given, then it might be appropriate to consider the results (speech analysis) given the propositions that 'The speaker is Mr B' and 'The speaker is an unknown person from the same region'.* By indicating that the '*observations of the police officers were uncontested*', we underline that these observations do not need to be assessed by a forensic scientist and thus are not part of 'E'.

If now, these observations are contested *(i.e., they are not a given)* and thus one asks an expert to assess this result*, then propositions such as 'The speaker is Mr B' and 'The speaker is an unknown person' would be more adequate.* Does this mean that one cannot use the technical system advocated by Morrison et al. [1] for voice comparison and evaluation? Of course not. The solution lies in the definition of the evidence. Morrison et al. [1] already offer part of this solution by stating that the forensic scientist could first assess the strength of the evidence of the accent based on the demographic information, followed by an assessment of the strength of the evidence based on the analysis of the acoustic properties of the speech within the population of people with the same accent.

One can consider that the evidence (E) is made of two aspects: the accent (E1) and the acoustic properties (E2). These aspects can be assessed by the same person, if s/he has expertise in both areas, or by two persons. One will assess E1 given the propositions that *'The speaker is Mr B'* and *'The speaker is an unknown person'*, thus Pr(E1|H). One will then assess E2 given the same propositions and given that one had observed a particular accent (E1), thus Pr(E2|H, E1). We see that the definition of the evidence in two parts allows the assessment of all the contested observations (both the accent and the acoustic properties), without mixing propositions and evidence. Finally both aspects of the evidence are combined, resulting in one likelihood ratio for the proposed propositions.

If one does not have the information on whether or not the observations on the accent are contested, then one can disclose this element and say that the evaluation is based on the assumption that the relevant population is a group of people from a given region.

Morrison et al. state:

Whether the submission of the recordings was filtered by a non-expert listener or not, it will usually be obvious to a forensic speech scientist whether the questioned speaker is male or female, what language they are speaking, and broadly what accent they are speaking. These properties will usually also be perceptually salient to all parties (prosecution, defence, judge, and trier of fact), either immediately or after they have been pointed out by the forensic speech scientist. If this is the case, speaker sex, and language and accent spoken can be taken as background information I, not as evidence E. They can therefore be used to refine the relevant population specified in the defence hypothesis.

Conducting an analysis based on a relevant population refined in this way will actually help avoid double counting.

We do not agree with this standpoint. Speaker sex, language and accent are all part of the evidence unless the case information dictates otherwise. All properties, even those obvious to the general public have an evidential value which needs to be assessed. It is up to the expert to determine if s/he is better capable of determining the evidential value of these features than the court and if so choose propositions allowing the evaluation of mentioned features.

As for the danger of double counting of the evidence, this may easily be avoided by explicitly mentioning which features have been accounted for in the expert's evaluation.

Even though we disagree with Morrison et al on this issue, we wish to thank them for stimulating further discussion and giving us the opportunity to remind the readership that it is helpful to take time for the definition of what is the evidence and what are the propositions and information. Forensic scientists will assess the former and the Court the latter. It is essential to acknowledge

this element when helping formulate propositions (that should reflect defence and prosecution positions without knowing the evidence which is to be evaluated).

To conclude, the tone of the remarks by Morrison et al invites us to make a general reflection on scientific debate. The authors seem to convey at numerous places in their text the view that only experts with extensive casework experience are qualified to debate the interpretative issues associated with their field of expertise. We disagree with that position: an open scientific debate, in an atmosphere of mutual respect, is a key enabler to progress, especially when it comes to the complexity of interpretative issues in forensic science. In 1998, at the RLA2C meeting in Avignon, a team - constituted of one of us and of a researcher in speaker recognition - presented a paper, later published in 2000 [4]. Today it is recognised by many experts in the field that this paper paved the way towards adopting a likelihood ratio approach in automatic speaker recognition. That paper is the product of a fruitful and constructive exchange of ideas between two researchers coming from different forensic specialities. We firmly believe that no progress will be achieved in forensic science if forensic scientists continue to exclude or diminish contributions from outsiders (sometimes called a "bunker mentality") and celebrate specialisms without promoting an interdisciplinary debate characterised by mutual respect. The paper by Meuwly and Champod [4] prefigured our vision of developing a forensic science culture: we still believe that this is the way forward.

Acknowledgements

References

[1] G.S. Morrison, Enzinger E., Zhang C., Letter to the Editor: Refining the relevant population in forensic voice comparison – A response to Hicks et alii (2015) The importance of distinguishing information from evidence/observations when formulating propositions, Sci. Justice 56 (2016) 492–497, http://dx.doi.org/10.1016/j.scijus.2016.07.002.

[2] T. Hicks, A. Biedermann, J.A. de Koeijer, F. Taroni, C. Champod, I.W. Evett, The importance of distinguishing information from evidence/observations when formulating propositions, Sci. Justice 55 (2015) 520–525, http://dx.doi.org/10.1016/j.scijus. 2015.06.008.

[3] A. Biedermann, T. Hicks, R. Voisard, F. Taroni, C. Champod, C. Aitken, I. Evett, E-learning initiatives in forensic interpretation: report on experiences from current projects and outlook, Forensic Science International 230 (2013) 2-7, http://dx.doi.org/10.1016/j.forsciint.2012.10.011.

[4] C. Champod, D. Meuwly, The Inference of Identity in Forensic Speaker Recognition, Speech Communication, 31 (2000) 193-203, http://dx.doi.org/10.1016/S0167-6393(99)00078-3