

Commentary on: I.E. Dror, G. Langenburg, “Cannot Decide”: The Fine Line Between Appropriate Inconclusive Determinations Versus Unjustifiably Deciding Not To Decide, *Journal of Forensic Sciences*, doi: 10.1111/1556-4029.13854 (in press)

Alex Biedermann^{1,*}, Joëlle Vuille¹, Silvia Bozza^{2,1}, Franco Taroni¹

¹University of Lausanne, Faculty of Law, Criminal Justice and Public Administration
School of Criminal Justice, 1015 Lausanne-Dorigny, Switzerland

²Ca’ Foscari University of Venice, Department of Economics, 30121 Venice, Italy

A. Biedermann, J. Vuille, S. Bozza, F. Taroni, Commentary on: I.E. Dror, G. Langenburg, “Cannot Decide”: The Fine Line Between Appropriate Inconclusive Determinations Versus Unjustifiably Deciding Not To Decide, *Journal of Forensic Sciences*, doi: 10.1111/1556-4029.13854, Epub 2018 Jul 5.

Dear Editor,

In their recent paper, Dror and Langenburg investigate “(...) under which circumstances inconclusive decisions are justified and even warranted (...), and, conversely, under what circumstances inconclusive decisions are unjustifiable and should not be permitted” [9, p. 1].¹ The authors approach these questions from the perspective of a “cognitive model which takes into account that decisions are an outcome of interactions and intersections between the actual data and human cognition” [9, p. 1]. In their analysis, the authors challenge selected aspects of inconclusive decisions on grounds that we believe have the potential to generate controversy. For example, they consider that some inconclusive decisions are “broad and imprecise” [9, p. 5]. The authors also contend that with inconclusive decisions, in some situations, “(...) the weight of the evidence may not be properly conveyed” [9, at p. 5].

While we welcome the authors’ perspective of considering expert conclusions as decisions, we do not share their overall negative tone with regards to inconclusive decisions, as well as some of the key arguments they present in support of their position. They offer grounds for a wider discussion on inconclusive determinations in forensic science. In this commentary, we discuss a number of claims that are central to Dror and Langenburg’s argument and evaluate them in light of a formal approach to decision analysis. We will show that the authors’ assertions, although *descriptively* plausible, not only fall short of a *normative* and *analytical* account of forensic identification conclusions, but also clash with general principles of interpreting forensic science evidence. Thus, the authors’ views remain unsubstantiated with respect to some of the ways, especially the formal ones, in which forensic identification conclusions can be conceptualised and understood. Below, we will discuss 6 points in turn.

1. A recurrent assertion by Dror and Langenburg is that an examiner’s conclusion should reflect the amount of available evidence. For example, they contend that

- “(...) an inconclusive decision should be a reflection of the limited weight of evidence” [9, p. 3] and
- “(...) if there are sufficient data and information in the evidence, then one can, and indeed must, make a decision regarding the source of the latent print (...) – put more bluntly, one cannot, and should not, make an inconclusive decision” [9, p. 4].

In essence, the authors thus maintain that if there is a lot of information, one should tend towards the conclusion ‘identification’ (or ‘exclusion’); if there is limited evidence, on the other hand, an ‘inconclusive’ decision is more appropriate and should be preferred. This amounts to defining the decision criterion as a function of the informative content ascribed to the evidence.² This idea is – from our point of view – short-sighted in two respects, however.

* *Corresponding author:* alex.biedermann@unil.ch; Visiting Researcher at University of Michigan Law School, Ann Arbor (MI, US).

¹ We refer to the generic page numbering of paper in press.

² By ‘evidence’ we mean here scientific evidence only, in particular the results of the comparison between a crime scene mark and a reference print from a person of interest, and *not all* evidence, scientific and otherwise, that may have a bearing on whether or not the person of interest is the source of the crime mark.

- First, if the suggestion here is that the value of evidence directly maps into probabilities for the propositions that the person of interest or an unknown person, respectively, is the source of the mark, then probability theory tells us that this is not so: whatever the value of the evidence, and the proposition it favours compared to a given alternative, *nothing* can be said about the probability of those propositions without specifying what those probabilities are before considering the evidence of interest. For example, although the probative value of a comparison between a crime scene mark and a reference print from a person of interest may be high (and in favour of the proposition of common source), nothing allows one to assert that the probability that the person of interest is actually the source of the mark is high, too – except one makes suitable assumptions about the prior probability for the competing propositions of interest. Indeed, if the prior probability for the proposition of common source was (very) small (before considering the evidence), it may remain only moderate or even low, *despite* a high probative value of the fingerprint comparison (favouring the proposition of common source).
- Second, even though one may reach a high probability that the person of interest is the source of the crime scene mark, one is still a step away from rendering an identification conclusion. Most importantly, just as probative value by itself says nothing about the probability of common source (see above), the probability for the proposition of common source does not by itself inform one about the conclusion that should be reached. In particular, even though one’s probability for the proposition of common source may be high, this does not mean that the conclusion ‘identification’ is warranted or justified. What is more, an identification may even be perfectly sensible even though the proposition of common source is *not* high. The whole discussion misses an essential point: we do not only decide based on what we think is actually the case (i.e., who the source of the crime mark is), but also based on our perceived desirability/undesirability of decision consequences (i.e., preferences³). For example, we may prefer a correct identification (i.e., the consequence of deciding ‘identification’ if the person of interest truly is the source of the crime mark) to a false identification (i.e., the consequence of wrongly deciding ‘identification’ if the person of interest is *not* the source of the crime mark). We may also prefer a missed identification (i.e., not identifying a person of interest who is actually the source of the crime mark), even though being an inaccurate decision consequence, to a false identification. Now imagine, for the purpose of illustration, that we have a high aversion to false identifications, for whatever reason (e.g., policy matters),⁴ or we may even want to avoid false identifications altogether. Especially in the latter case, it will be sensible for us *not* to identify, despite a high probability for the proposition that the person of interest is the source of the crime mark and/or the fact that we assign a high probative value to the fingerprint comparison.⁵ These are not just intuitive considerations, but can be shown to follow from formal decision theoretic analyses that we have exposed elsewhere [e.g., 3, 4, 13].⁶

It follows from the above that the assertion that particular amounts of data (or evidential value) should lead to a particular conclusion (identification or exclusion), and hence discourage an ‘inconclusive’ conclusion, is not supported in a formal decision theoretic account of forensic identification. This also holds for other assertions such as “[a]n inconclusive decision should be a reflection of the limited weight of evidence” [9, p. 3], “inconclusive decisions do not represent sufficient weight of the evidence” [9, p. 3] and “[t]he inconclusive category (...) does not properly convey the estimated weight of evidence” [9, p. 4]. This is because there is more to decisions than just weight of evidence and also because decisions do *not* directly reflect weight of evidence (see also point no. 3 below).

2. The authors refer to our paper [3] “(...) for a discussion of the utility function wherein false positives and false negatives have negative utility, and inconclusive decisions have no utility” [9, p. 3]. This is incorrect for several reasons. First, we did not assign ‘negative utilities’ to decision consequences: we chose a (0, 1) scale (though nothing prevents one to use a different scale, including negative values). Second, we did not assign ‘no utility’ to inconclusive decisions. Let us emphasize that:

- It is important to understand that utilities are assigned to decision consequences, *not* to the decisions themselves – hence it is unsound to talk in terms of the ‘utility of a decision’. Decisions have *expected* utilities, but those are not assigned. They merely follow by definition from a computational step based on the utilities assigned to decision consequences and probabilities assigned to states of nature.
- In the analysis presented in [3] the *consequences* of inconclusive decisions have, as all other decision consequences, an assigned utility, though it may be equal to the minimum value zero. Actually, in our analyses, we assigned positive values to the utilities of the consequences of inconclusive decisions.

³Note that here we do not understand the term ‘preference’ in the same way as Dror and Langenburg in their Fig. 1 [9, p. 3].

⁴See [5] further discussion on such particular preferences.

⁵As an aside, note that Dror and Langenburg consider instances where “(...) the weight of evidence is so great that an “inconclusive” is inappropriate (...)” [9, p. 4] as ‘errors’. This does not follow from our arguments exposed here. See also our below point no. 4 for further discussion on the issue of error.

⁶It also follows from these considerations that in some cases, the preference system may be such that even in presence of a small amount of information, nothing will prevent one from reaching a conclusion in terms either of an identification or exclusion. This is because one may have either a high prior probability, and/or again because one maintains peculiar desirabilities/undesirabilities regarding decision consequences.

Hence, when interpreting the terms ‘decision’ and ‘utility’ in the logical sense of decision theory, statements such as “We must remember that inconclusive decisions have a net utility of zero (...)” [9, p. 3] find no justification from a normative point of view.

3. In situations in which the examiner is close to making an ‘identification’ conclusion, but cannot actually render this conclusion, Dror and Langenburg consider that:

“(...) inconclusive decisions entail a loss of valuable data. Because the decision is reported as “inconclusive,” the weight of the evidence is lost by the imprecision and breath of the vague “inconclusive” category.” [9, p. 4]

To see why this assertion can generate misunderstandings, consider again our arguments exposed in the above point no. 2, and recall that there is no direct exclusive transition from weight of evidence to decision. It is crucial to understand that weight of evidence, first and foremost, is concerned with inference. In our context of application, this means that the weight of evidence that is assigned to results of a fingerprint comparison will impact on one’s assessment of the probability that the person of interest, rather than an unknown person, is the source of the crime mark. Technically speaking, this leads to a posterior probability. As we have seen, however, this probability is only one preliminary, among others, to making a decision. So, we may decide whatever we want, our data will *not* be lost: it remains encapsulated in the posterior probability for the main propositions of interest. Stated otherwise, our probabilistic assessment of the proposition according to which the person of interest is the source of the crime mark, rather than an unknown person, is not impacted by the subsequent decisions that we make.⁷ Weight of evidence may only be said to be ‘lost’ if we do not use it for its intended purpose: i.e., if we do not use it to revise our beliefs about the propositions of interest.

4. A further issue that Dror and Langenburg raise is the question of “(...) when inconclusive decisions can be considered an error and when they are justified” [9, p. 4]. The point on which we wish to focus here is the use of the term ‘error’. Is it sensible to try to qualify inconclusive decisions as erroneous? Conventionally, ‘error’ is reserved for situations of factual errors: for example, if the examiner concludes ‘identification’ when in reality the person of interest is not the source of the crime scene mark, the examiner’s decision to identify is not accurate – it is an erroneous determination. The same holds for missed identifications. However, we can qualify examiners’ decisions in this way only if we know the underlying truth state (i.e., whether or not the person of interest is truly the source of the crime mark), which may be the case in experiments under controlled conditions (such as proficiency tests), but rarely ever⁸ in operational casework. It may be questioned, thus, whether it is worthwhile to use the term ‘error’ for inconclusive decisions. As per definition, inconclusive decisions do not make a ‘call’ in one way or the other regarding the proposition of whether or not the person of interest is the source of the crime mark, hence the term ‘error’ is to no effect, even in experiments under controlled conditions where the underlying truth-state is known. Yet another possibility of considering error is non-conformity with procedures and predefined rules (of decision), which might go into the direction of the meaning intended by Dror and Langenburg; but still, this option is not workable since the authors do not provide any clear rule of decision, except the tentative suggestion of relating it to value of evidence (see also our discussion points no. 1 and 3 above).

5. As regards future developments, Dror and Langenburg foresee that “[t]he ability to appropriately estimate the weight of the evidence and then communicate that to the triers of fact goes a long way toward the issue of imprecise categorical decisions. The vague, broad category of inconclusive will become a discarded artifact of the past if and when a proper and validated model is employed” [9, p. 4]. While we agree that the availability of appropriate measures for weight of evidence, which we take to be likelihood ratios, would make identification decisions superfluous for forensic scientists – because they would then report likelihood ratios – identification conclusions as such will not necessarily disappear. Recipients of expert information may still need to make decisions regarding common source, once they have taken into account the weight of evidence as reported by forensic scientists. In particular, a court will always be able to say that an item of evidence did not help them at all in the case, either way: they may take a decision regarding the ultimate issue but not taking a stance with regards to the question of common source, hence considering it ‘inconclusive’.

6. In their conclusions, Dror and Langenburg state:

“Understanding forensic (and other expert) decision-making requires to understand the nature of the data, but also the human cognitive processes involved in decision-making” [9, p. 5] and
“Our model (...) illustrated the importance of taking into account the underlying cognitive processes involved in decision-making.” [9, p. 5]

⁷Note that this results from an assumption of non-conditioning of probabilities by decisions, which is suitable for applications in forensic inference. There are exceptions to this in other areas of application (e.g., medicine) where, for example, the probability of the future state of health depends on the type of treatment chosen (e.g., surgery, non-invasive treatments, etc.). See also discussion in [11].

⁸Examples for exceptions are cases in which the circumstances and all other elements are such that the expert’s conclusion is *judged* erroneous (see, e.g., the Mayfield case).

We do not question that it is important to inquire about how forensic scientists arrive at their conclusions and develop theories that describe the thinking and the patterns of decision behaviour of scientists. This helps recognise circumstances and factors that favour, respectively that endanger, accurate decision-making. But, at the end of the day, the primary questions of interest are ‘how to think *logically* about the problem of forensic identification?’ and ‘what are the fundamental logical ingredients of forensic identification decisions?’. As we have argued previously [e.g., 5, 6],⁹ these questions are of a normative nature and should be distinguished from a descriptive perspective. By normative we mean here a logical reference point against which observable thinking and decision behaviours can be compared. To take a simple example borrowed from De Finetti [8] and Lindley [12], imagine we wish to inquire about how to reasonably combine numbers, such as $2 + 2$. The normative approach here consists in the arithmetic operation of addition which directs us to solve the problem logically. An empirical and descriptive approach might focus on asking people their opinion, with the possibility of encountering answers different from 4. While it may be helpful to know whether people deviate from the arithmetically correct results, and under what circumstances, that inquiry is pointless without a normative theory available in the first place, against which observable behaviour can be compared. To take a forensic example, consider the well known prosecutor’s and defence attorney’s fallacies [14]. These are important insights about *observable* patterns of reasoning under uncertainty; but their relevance only comes to light because of the underlying normative account (i.e., the probabilistically ‘right’ result) against which people’s intuitive reasoning is evaluated. So, what we really need are fundamental understandings of how to logically make up our minds and of what questions are relevant to ask. For the illustrative problem of $2 + 2$, people can improve by learning about arithmetics. The mere observational result that uninstructed people’s answers may spread over some range, though adding to phenomenological knowledge, is not helpful to advance fundamental understandings.¹⁰ For the case of reasoning under uncertainty (inference), we need people to learn about the logic of the probably (i.e., probability theory). Arguably, for forensic decision making under uncertainty we need a logical extension of probabilistic reasoning, which is given by decision theory. This conclusion, however, is not prescriptive: the principles of decision logic as such do not contain any binding elements [7] – as much as there is nothing in probability theory to say that we *must* apply it. However, the stakes involved in forensic reporting are such that we do not see what else than (decision) logic should be considered acceptable.

Acknowledgments

Alex Biedermann acknowledges the support provided by the Swiss National Science Foundation through Grant No. BSSGI0.155809 and the University of Michigan Law School, Ann Arbor, MI (Michigan Grotius Research Fellowship).

References

- [1] J. Baron. *Thinking and Deciding*. Cambridge University Press, New York, 4th edition, 2008.
- [2] J. Baron. The point of normative models in judgment and decision making. *Frontiers in Psychology*, 3, Article 577:1–3, 2012.
- [3] A. Biedermann, S. Bozza, and F. Taroni. Decision theoretic properties of forensic identification: underlying logic and argumentative implications. *Forensic Science International*, 177:120–132, 2008.
- [4] A. Biedermann, S. Bozza, and F. Taroni. The decisionalization of individualization. *Forensic Science International*, 266:29–38, 2016.
- [5] A. Biedermann, S. Bozza, F. Taroni, and P. Garbolino. A formal approach to qualifying and quantifying the ‘goodness’ of forensic identification decisions. *Law, Probability and Risk*, page in press, 2018 (in press).
- [6] A. Biedermann, F. Taroni, and C. Aitken. Liberties and constraints of the normative approach to evaluation and decision in forensic science: a discussion towards overcoming some common misconceptions. *Law, Probability and Risk*, 13:181–191, 2014.
- [7] B. de Finetti. Dans quel sens la théorie de la décision est-elle et doit-elle être “normative”. In *Colloques Internationaux du Centre National de la Recherche Scientifique, “La Decision”*, pages 159–169, Paris, 1961. Éditions du Centre National de la Recherche Scientifique.
- [8] B. de Finetti. *Theory of Probability, A Critical Introductory Treatment, Volume 1*. John Wiley & Sons, London, 1974.

⁹For a discussion in the context of decisions other than identification, arising with fingerprints, see e.g. [10].

¹⁰Note that we are not dismissive of descriptive theories, as much as Dror and Langenburg are not dismissive of normative concepts such as probability and utility. We are only arguing here that in forensic science descriptive approaches take value in the light of a normative account, and that the full scope of decision making analyses needs to be kept in mind, covering normative, descriptive and prescriptive perspectives [1, 2].

- [9] I. E. Dror and G. Langenburg. “Cannot decide”: The fine line between appropriate inconclusive determinations versus unjustifiably deciding not to decide. *Journal of Forensic Sciences*, in press:1–6, 2018.
- [10] S. Gittelsohn, S. Bozza, A. Biedermann, and F. Taroni. Decision-theoretic reflections on processing a fingerprint. *Forensic Science International*, 226:e42–e47, 2013.
- [11] D. V. Lindley. *Making Decisions*. John Wiley & Sons, Chichester, second edition, 1985.
- [12] D. V. Lindley. *Understanding Uncertainty*. John Wiley & Sons, Hoboken, 2006.
- [13] F. Taroni, S. Bozza, A. Biedermann, G. Garbolino, and C. G. G. Aitken. *Data Analysis in Forensic Science: a Bayesian Decision Perspective*. Statistics in Practice. John Wiley & Sons, Chichester, 2010.
- [14] W. C. Thompson and E. L. Schumann. Interpretation of statistical evidence in criminal trials: The prosecutor’s fallacy and the defense attorney’s fallacy. *Law and Human Behaviour*, 11:167–187, 1987.