Science and Justice 55 (2015) 520–525

Contents lists available at ScienceDirect

Science and Justice

journal homepage: www.elsevier.com/locate/scijus

The importance of distinguishing information from evidence/observations when formulating propositions

T. Hicks [a,b,*], A. Biedermann [a,b], J.A. de Koeijer [c], F. Taroni [a], C. Champod [a], I.W. Evett [d]

[a] School of Criminal Justice, University of Lausanne, 1015 Dorigny, Switzerland
[b] Fondation pour la formation continue universitaire lausannoise (UNIL-EPFL), 1015 Dorigny, Switzerland
[c] Netherlands Forensic Institute, NFI, PO Box 24044, 2490 AA, The Hague, The Netherlands
[d] Principal Forensic Services Ltd, Office 4, 219 Kensington High Street, London W8 6BD, United Kingdom

ARTICLE INFO

Article history:
Received 15 May 2015
Received in revised form 19 June 2015
Accepted 25 June 2015

Keywords:
Propositions
Interpretation
Likelihood ratio
Evidence
Information
Evaluation

ABSTRACT

The value of forensic results crucially depends on the propositions and the information under which they are evaluated. For example, if a full single DNA profile for a contemporary marker system matching the profile of Mr A is assessed, given the propositions that the DNA came from Mr A and given it came from an unknown person, the strength of evidence can be overwhelming (e.g., in the order of a billion). In contrast, if we assess the same result given that the DNA came from Mr A and given it came from his twin brother (i.e., a person with *the same DNA profile*), the strength of evidence will be 1, and therefore neutral, unhelpful and irrelevant[1] to the case at hand. While this understanding is probably uncontroversial and obvious to most, if not all practitioners dealing with DNA evidence, the practical precept of *not* specifying an alternative source with the same characteristics as the one considered under the first proposition may be much less clear in other circumstances. During discussions with colleagues and trainees, cases have come to our attention where forensic scientists have difficulty with the formulation of propositions. It is particularly common to observe that results (e.g., observations) are included in the propositions, whereas—as argued throughout this note—they should not be. A typical example could be a case where a shoe-mark with a logo and the general pattern characteristics of a Nike Air Jordan shoe is found at the scene of a crime. A Nike Air Jordan shoe is then seized at Mr A's house and control prints of this shoe compared to the mark. The results (e.g., a trace with this general pattern and acquired characteristics corresponding to the sole of Mr A's shoe) are then evaluated given the propositions 'The mark was left by Mr A's Nike Air Jordan shoe-sole' and 'The mark was left by an unknown Nike Air Jordan shoe'. As a consequence, the footwear examiner will not evaluate part of the observations (i.e., the mark presents the general pattern of a Nike Air Jordan) whereas they can be highly informative. Such examples can be found in all forensic disciplines.

In this article, we present a few such examples and discuss aspects that will help forensic scientists with the formulation of propositions. In particular, we emphasise on the usefulness of notation to distinguish results that forensic scientists should evaluate from case information that the Court will evaluate.

1. Introduction

The value of observations depends crucially on propositions and information. Being rigorous and logical when defining observations and propositions is not only of paramount importance for the Court (so that it is known how and what findings were evaluated), but it also helps scientists to identify the issue they can help with, to ask the right questions and thus give case tailored answers.

Because evaluation of forensic observations (e.g., evidence) needs to take place within a logical structure it is useful to remind the reader of the principles of interpretation suggested for example in Evett et al. [3]. These three principles can be reconstructed as features that conform to Bayes' theorem, and hence qualify as logical.

The first principle of interpretation emphasises that evaluation takes place within a framework of circumstances. This point is essential for two reasons: first, it is only from the circumstances that one can understand what is the issue that forensic science can help with in the given case, and what propositions need to be considered. Second, it underlines that probabilities are conditional and depend on what we know, what we are told and what we assume. Saying that they are conditional on the information and the

* Corresponding author at: Ecole des sciences criminelles, School of Criminal Justice, Bâtiment Batochime, 1015 Lausanne-Dorigny, Switzerland. Tel.: +41 21 692 46 07; fax: +41 21 692 46 05.
E-mail address: Nathalie.HicksChampod@unil.ch (T. Hicks).

[1] Relevant evidence is defined here as any evidence that has the potential to alter one's view of the competing propositions of interest [1]. This use of the term 'relevance' is to be distinguished from 'relevance' as a descriptor of a trace's pertinence, that is the probability that it has been left by the offender [2].

http://dx.doi.org/10.1016/j.scijus.2015.06.008
1355-0306/

knowledge that we have means also that they are personal and case specific.

The second principle of forensic interpretation says that forensic results ought to be evaluated given at least two competing propositions. These propositions are generally referred to as reflecting the viewpoints of, respectively, prosecution and defence.

The third principle clarifies the role of the expert and the role of the legal decision maker. It outlines that forensic scientists ought to give their assessment on the results and not on the propositions. It is for the Court to assess propositions. This last principle has been discussed in both literature and practice for numerous years, but it appears to cause continuing trouble among participants in legal proceedings. Indeed, it is still quite common to see scientists, the Court or the press, confuse the probability of the results with the probability of the propositions (thus transposing the conditional [3,4]).

For the discussion in this paper, it is not necessary to state Bayes' theorem in full detail. It is sufficient to define notation for some of its constituting elements. Prosecution and defence propositions will be denoted respectively H_p and H_d. I will stand for the information, impacting the value of the findings. The observations to be evaluated will be denoted E. The notation used for probability will be $Pr(\ |\ )$ where the vertical line is a conditioning bar and everything that appears to the right of it is taken as conditional information.

Below, we will present and discuss several examples of cases where forensic practitioners expressed difficulties in the formulation of propositions. First, we will show an example where results appear in propositions (which should not happen), and then some examples where part of the observations are taken as information and included in the propositions.

As an aside, let us anticipate that if results are included in the propositions, then this means that it will be for the Court to evaluate these results, without resorting to any advice from the scientist. That is, since the Court needs to hold an initial opinion about propositions of interest (i.e., a prior probability) before considering the forensic observations, propositions that include a statement about analytical features will require the Court to make an *informed* initial judgment, hence an assessment of those features. This is acceptable if no forensic knowledge is required to do so, but problematic if the scientist's duty is left to the Court. This consequence of the definition of propositions is important for practical proceedings and will be one point of discussion in the examples outlined below.

2. Case examples

We have emphasised that forensic observations need to be distinguished from information that will be assessed by the Court. This is crucial given two conditions: first, if expert knowledge is needed to evaluate the findings and secondly if the observations have value (i.e., they allow discrimination, that is, they are relevant as defined above). Evett et al. [5] have discussed criteria that will help spotting whether propositions are formulated in a useful way:

1. Propositions come in pairs.
2. Propositions need background information.
3. Propositions are formal.
4. Propositions are mutually exclusive and exhaustive in the context of the case[2].
5. Prior odds relate to propositions.
6. Propositions relate to inference.

To those criteria, we would like to add the following guides:

7. Propositions are about ‘causes’ (i.e., target events that lead to particular traces or findings), not results. Forensic observations should not be included in propositions. Although, if these observations have no value (i.e., they allow no discrimination) or if their value can be assigned without forensic knowledge, then including these observations in propositions will have no impact and is thus acceptable.
8. Propositions should not be findings led: to ensure it is the case, the formulation of propositions should be made *before* comparing the trace to a potential source.

2.1. Mixing results and propositions

2.1.1. Example 1 (DNA case)

Let us take a case typically used as an exercise with trainees. Imagine that there has been a burglary in your town. The police are called to the scene of crime. They collect a DNA trace from Mr Smith's safe that has been forced open. The trace is submitted to the laboratory. A suspect, Mr B. is arrested and his DNA profile is analysed and compared to that of the trace. He says he has nothing to do with the incident and has never been in Mr Smith's house. Experience shows that many practising scientists would suggest or be happy to accept propositions of the following kind: ‘The matching DNA comes from Mr B.’ versus ‘The matching DNA comes from some unknown unrelated person’. Is there a problem with these sub-source level propositions [3]? If we use some further notation, we can see that there clearly is, as shown below:

H_p	The matching DNA comes from Mr B.
H_d	The matching DNA comes from an unknown person[3].
I	The crime took place in your town.
E	The DNA profile of the trace from the crime scene matches Mr B.

If we evaluate the observations (i.e., E: the DNA profile derived from the trace matches Mr B's DNA profile.) given the proposition that ‘The matching DNA comes from Mr B’, then clearly results appear in the proposition. This will be most evident when assigning the denominator. In fact, if defence's proposition states that an unknown person has left matching DNA, the profile of the trace will obviously match Mr B. Hence, $Pr(E|H_d,I) = 1$, as is $Pr(E|H_p,I)$. In that case, the result E does not provide any guidance as to which of the two propositions is true. This example illustrates that if we want to assess the value of the results meaningfully, then they should not overlap with the propositions. An example of useful definition of propositions, information and observations would be:

H_p	The DNA comes from Mr B.
H_d	The DNA comes from an unknown person.
I	The crime took place in your town.
E	The trace presents an unambiguous single DNA profile denoted G_C; Mr B's DNA profile is denoted G_B; $G_B = G_C$

Although the ‘contamination’ of propositions with results is an issue in this example, there is perhaps a more fundamental problem with the oversimplifying word ‘match’, that now is almost abusively used across

[2] This means that the propositions of defence and prosecution cannot both be true and that there are no other propositions that -at the time of the evaluation- appear to be relevant to the case. Both these conditions have to be met in order for the evaluation to be coherent.

[3] Sometimes propositions and observations are verbally blended so that it becomes obscure under what proposition the results are being considered. An example is the formulation ‘Probability of a coincidental match’. This expression suggests a focus on the observations in the event that the suspect is not the source, that is a negation of the first proposition. It is well known, however, that the simple negation of the first propositions rarely provides a viable alternative because, if the suspect is not the source, necessarily someone else is, and it will become relevant to enquire about whom this someone else is (e.g., a member of the general population, a closed set of possible offenders or a family member).