Science and Justice 56 (2016) 364–370

Contents lists available at ScienceDirect

Science and Justice

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

Discussion on how to implement a verbal scale in a forensic laboratory: Benefits, pitfalls and suggestions to avoid misunderstandings

Raymond Marquis *, Alex Biedermann, Liv Cadola, Christophe Champod, Line Gueissaz, Geneviève Massonnet, Williams David Mazzella, Franco Taroni, Tacha Hicks

School of Criminal Justice, University of Lausanne, 1015 Dorigny, Switzerland

ARTICLE INFO

Article history:
Received 3 March 2016
Received in revised form 23 May 2016
Accepted 24 May 2016

Keywords:
Verbal scale
Interpretation
Likelihood ratio
Probative value
Statement

ABSTRACT

In a recently published guideline for evaluative reporting in forensic science, the European Network of Forensic Science Institutes (ENFSI) recommended the use of the likelihood ratio for the measurement of the value of forensic results. As a device to communicate the probative value of the results, the ENFSI guideline mentions the possibility to define and use a verbal scale, which should be unified within a forensic institution. This paper summarizes discussions held between scientists of our institution to develop and implement such a verbal scale. It intends to contribute to general discussions likely to be faced by any forensic institution that engages in continuous monitoring and improving of their evaluation and reporting format. We first present published arguments in favour of the use of such verbal qualifiers. We emphasise that verbal qualifiers do not replace the use of numbers to evaluate forensic findings, but are useful to communicate the probative value, since the weight of evidence in terms of likelihood ratio are still apprehended with difficulty by both the forensic scientists, especially in the absence of hard data, and the recipient of information. We further present arguments that support the development of the verbal scale that we propose. Recognising the limits of the use of such a verbal scale, we then discuss its disadvantages: it may lead to the spurious view according to which the value of the observations made in a given case is relative to other cases. Verbal qualifiers are also prone to misunderstandings and cannot be coherently combined with other evidence. We therefore recommend not using the verbal qualifier alone in a written statement. While scientists should only report on the probability of the findings – and not on the probability of the propositions, which are the duty of the Court – we suggest showing examples to let the recipient of information understand how the scientific evidence affects the probabilities of the propositions. To avoid misunderstandings, we also advise to mention in the statement what the results do not mean. Finally, we are of the opinion that if experts were able to coherently articulate numbers, and if recipients of information could properly handle such numbers, then verbal qualifiers could be abandoned completely. At that time, numerical expressions of probative value will be appropriately understood, as other numerical measures that most of us understand without the need of any further explanation, such as expressions for length or temperature.

1. Introduction

In a recently published guideline for evaluative reporting[1] in forensic science, the European Network of Forensic Science Institutes [2] recommended the use of the likelihood ratio for the measurement of the value of forensic results.[2] The document specifies a series of principles to guide the scientist's thinking about the evaluation of forensic results: the first principle is that interpretation takes place in a framework of circumstances and that the value of forensic observations depends on relevant case information. In other words, if the case information changes, the interpretation of the findings must be reviewed as well. The second principle states that the forensic observations shall be interpreted

* Corresponding author.
E-mail addresses: raymond.marquis@unil.ch (R. Marquis), alex.biedermann@unil.ch (A. Biedermann), liv.cadola@unil.ch (L. Cadola), christophe.champod@unil.ch (C. Champod), line.gueissaz@unil.ch (L. Gueissaz), genevieve.massonnet@unil.ch (G. Massonnet), williams.mazzella@unil.ch (W.D. Mazzella), franco.taroni@unil.ch (F. Taroni), nathalie.hickschampod@unil.ch (T. Hicks).

[1] Please note that our discussions will neither include investigative nor technical reporting as defined in both the statement of the Association of Forensic Science Providers [1] and the recent ENFSI Guideline [2].

[2] A likelihood ratio is a ratio of two probabilities: the probability of the observations given that the first proposition and the conditioning information are true, divided by the probability of the observations given that the alternative proposition and the conditioning information are true.

http://dx.doi.org/10.1016/j.scijus.2016.05.009
1355-0306/

in light of at least one pair of competing propositions. The third principle stresses that it is appropriate for the scientist only to address the probability of the observations given the propositions, and not the probability of the propositions themselves [3,4,5].

The application of these principles ensures that the approach is balanced and logical. In particular, with the value of the likelihood ratio, scientists express an opinion on the observations they have made and they convey the degree of support provided by these results for one proposition over the alternative. The recipients of expert information (e.g., the Court) can then use this information to update their belief on the competing propositions, considering all the other elements of the case.

Although the approach is well structured, its practical implementation represents a challenge to both individual scientists and forensic science institutions as a whole. The introduction of such a change in evaluation and reporting practice does not happen overnight and requires an institutional strategy over a period of time.[3]

As a device to support scientists' reporting practice, the ENFSI guideline mentions the possibility to define and use standardised verbal qualifiers for ranges of likelihood ratio values — also often referred to as a 'verbal scale'. There is no binding recommendation in the ENFSI guideline, but it is advised to use a single and unified reporting convention for all disciplines within a forensic institution. It must be emphasised that a verbal scale is not a replacement for the likelihood ratio value, but it can represent a convenient way to communicate this value.

This article reports on this particular aspect of forensic science reporting. Specifically, we will focus on questions and discussions that emerged from works towards the development and implementation of optional verbal qualifiers for probative value. This paper — organised in a question-answer format — intends to contribute to general discussions likely to be faced by any forensic institution that engages in continuous monitoring and improving of their evaluation and reporting format. We hope that our readers will see merit in this initiative, as we believe that sharing practical experience from different institutions regarding challenges, approaches and strategies for implementing the principles emphasised in the ENFSI guideline is essential.

2. Discussion of key questions

In the following sections, we describe the key elements raised in the discussions held within our institution. As it appears that verbal scales to communicate the weight of evidence are quite commonly used in the forensic community, we highlight in Section 2.1 the arguments in favour of such a practice. We then address, in Section 2.2, the question of whether such a verbal scale remains useful in the current state of practice. As it emerges that a verbal scale may still be of help today, Section 2.3 discusses the question of what type of verbal scale should be used. Section 2.4 covers misunderstandings that occur when expert opinion is only conveyed by words. To overcome these problems, one possible solution could be to add the full range of the verbal scale in the statement. Section 2.5 provides arguments against this suggestion. Finally, Section 2.6 proposes some recommendations that can help communicating the value of evidence.

2.1. What are the published arguments in favour of the use of a verbal scale as suggested in the ENFSI guideline for evaluative reporting in forensic science?

Harmonisation of conclusions and the use of common terminology are an important aspect in forensic science reporting. As early as 1979, Kind et al. [6] suggested conventions regarding categorisation of "*samples*" submitted to forensic science laboratories for examination. In 1986, Brown and Cropp [7] discussed the importance of avoiding some terms in reports and even went a step further by suggesting harmonisation of conclusions using a correspondence table between probabilities and adverbs. A year later, Leung and Cheung [8] commented on the wide range of terminologies used to define qualified opinions, and proposed a way to provide uniformity in this respect. One can see therefore that standardisation of qualitative terms to report the opinion of forensic scientists regarding the value of evidence has been a preoccupation for forensic scientists for a long time.

Authors focused on the choice of terms used to conclude and tried to find appropriate words to convey the value of the evidence in a case, but to our knowledge it was Evett [9] who first suggested adopting a verbal scale in forensic science in 1987, based on Jeffreys's book 'Theory of Probability', first published in 1939. We summarize hereafter arguments presented in favour of the use of a verbal scale.

Early arguments in favour of the use of a so-called 'verbal scale' are that it should promote logical reporting. Indeed, according to Evett [9], verbal qualifiers can help scientists express themselves on the value of the results given the propositions rather than on propositions themselves. By saying "the results [strongly] support defence proposition rather than prosecution proposition", scientists ensure that they do not transpose the conditional. The forensic observations are thus evaluated in agreement with a logical approach, and the scientist's statement of the value of the observations serves the purpose to assist the recipient of expert information going from prior to posterior odds [3].

A further argument in favour of the use of verbal equivalents is that in fields where structured, documented and published data are scarce, reporting a likelihood ratio may be a challenge. Indeed, some scientists feel that they can only commit themselves to particular numbers if they can trace them back to calculations based on hard numerical data [10,11]. Also it may give the illusion of a level of mathematical precision, whereas no calculation was carried out per se, but only an assignment based on experience and training. Instead, in those situations, scientists are generally more at ease by communicating the magnitude of their likelihood ratio with a verbal equivalent. For example, Jackson [10] wrote that "the scientist should evaluate broadly the magnitude for the likelihood ratio and translate that into a verbal equivalent" (p. 85).

Verbal conventions also cover an important third function, as one assumes that it helps communication between scientists and non-scientists. Indeed, faced with uncertainty, it is known that generally people prefer words to numbers [12] so that the use of verbal equivalents is considered to be beneficial for the recipients of expert information who do "not feel confident in handling numbers and react negatively to mathematical formulae" (p. 447) [13].

For the above reasons, it would seem convenient for scientists to use a verbal equivalent in their statement, instead of reporting the (numerical) value of the likelihood ratio alone: this verbal term would then express the value of the observations as given by the likelihood ratio. But, as we will see in the next sections, such a device for supporting reporting has some disadvantages that one needs to consider.

2.2. Do we need verbal equivalents in the current state of practice?

One negative aspect of verbal qualifiers is that scientists may fail to remember that "the assessment comes first, the decision about a verbal qualifier comes later" (p. 47) [14]. Ideally, as discussed by Berger et al. [14], the assignment of a numerical value, that is a probability (or probability density), is "preferable whenever possible" (p. 47). It is true that likelihood ratios can be assigned qualitatively, but experts need to express themselves in numerical terms as it is through numbers that we measure things, convey information and combine it with other information [12]. Without numbers, one cannot combine different types of traces if needed. One cannot really appreciate either the impact of the

[3] For this purpose, the ENFSI guideline proposes a general four step roadmap to help quality managers and leading scientists design an implementation plan that is flexible enough to be adapted to service specific requirements and needs.