Article

THE INTERNATIONAL JOURNAL of EVIDENCE & PROOF

The International Journal of Evidence & Proof
2017, Vol. 21(1-2) 14–29

DOI: 10.1177/1365712716674796
journals.sagepub.com/home/epj
SAGE

Development of European standards for evaluative reporting in forensic science: The gap between intentions and perceptions

Alex Biedermann
School of Criminal Justice, University of Lausanne, Switzerland

Christophe Champod
School of Criminal Justice, University of Lausanne, Switzerland

Sheila Willis
Forensic Science Ireland, Dublin, Ireland

Abstract

Criminal justice authorities of EU countries currently engage in dialogue and action to build a common area of justice and to help increase the mutual trust in judicial systems across Europe. This includes, for example, the strengthening of procedural safeguards for citizens in criminal proceedings by promoting principles such as equality of arms. Improving the smooth functioning of judicial processes is also pursued by works of expert working groups in the field of forensic science, such as the working parties under the auspices of the European Network of Forensic Science Institutes (ENFSI). This network aims to share knowledge, exchange experiences and come to mutual agreements in matters concerning forensic science practice, among them the interpretation of results of forensic examinations. For example, through its Monopoly Programmes (financially supported by the European Commission), ENFSI has funded a series of projects that come under the general theme 'Strengthening the Evaluation of Forensic Results across Europe'. Although these initiatives reflect a strong commitment to mutual understanding on general principles of forensic interpretation, the development of standards for evaluation and reporting, including roadmaps for implementation within the ENFSI community, are fraught with conceptual and practical hurdles. In particular, experience through consultations with forensic science practitioners shows that there is a considerable gap between the intentions of a harmonised view on principles of forensic interpretation and

Corresponding author:
Alex Biedermann, School of Criminal Justice, University of Lausanne, Switzerland; Litigation Law Unit, Law School, University of Adelaide, Adelaide, SA, Australia (Visiting Researcher).
E-mail: alex.biedermann@unil.ch

the way in which works towards such common understanding are perceived in the community. In this paper, we will review and discuss several recurrently raised concerns. We acknowledge practical constraints such as limited resources for training and education, but we shall also argue that addressing topics in forensic interpretation *now* is of vital importance because forensic science continues to be challenged by proactive participants in the legal process that tend to become more demanding and less forgiving.

Keywords
European context, evaluative reporting, forensic evidence

Introduction

Forensic interpretation, that is, the evaluation of the probative value of scientific findings, is a topic that arises in various contexts and stages in the judicial process. As an integral part of practical proceedings, evaluative challenges primarily pertain to the particular circumstances of cases, but a closer inspection of the subject suggests that organisational and academic dimensions should also be observed. Below we raise these three areas in turn.

Imagine a case of breaking and entering and assault on an elderly woman in her home. At the point of entry, a large fresh bloodstain is recovered and delivered to the laboratory for DNA analysis. Combination of a presumptive test and appearance makes it safe to assume that the stain is blood. The same night, based on the description provided by the victim, the police arrest a man. A reference DNA swab has been taken from him. The suspect says that he has never been in the premises. At the crime scene, a weapon is also found. It is swabbed to recover and secure any biological material, including any cells left by the person who used it. Following laboratory analyses, two DNA profiles were detected, one corresponding to the victim, and the other corresponding to the DNA profile of the suspect. 'Is this good evidence?' is a question that may be found appealing in such a case. Alternatively, it might also be asked if one could conclude that the suspect is the source of the recovered DNA, or whether the suspect is the assailant. Such questions may be the result of the stupefying effect of learning that the DNA profiles correspond, paired with the commonly held belief that a report on corresponding DNA profiles *must* necessarily mean something. Discussants may also struggle with the fact that DNA profiles from different traces corresponding with the profile of the *same* person may have substantially different probative values depending, for example, on the nature of the staining and the position and condition in which it has been found.

For several reasons, it is not very helpful to attempt a reply to this questioning at this juncture. One reason is that further questions are prompted. For example, when asking 'Is it good evidence?', an immediate reaction is to ask: 'Evidence for what?' This suggests that, first and foremost, we ought to enquire about the actual issue in the case and the needs of the members of the criminal justice system. It might also be advisable to consider what the person of interest says. Clearly, a case in which the suspect asserts that the weapon is his, but it was stolen from him a month ago, is fundamentally different from a case in which he asserts that he has nothing to do with the weapon. In the former situation, the question of whether the recovered DNA profile *comes* from the person of interest, that is, a question at the so-called source level, may be of limited interest only (Taroni et al., 2013). This exemplifies that evaluating scientific findings in the light of relevant case information is a crucial requirement (Champod, 2014a; Evett and Weir, 1998; Willis, 2014).

Over the past few years, the importance of sound assessment of scientific results in the context of the case at hand has also been recognised by expert working groups. For example, the European Network of Forensic Science Institutes (ENFSI) issued the following statement in 2010:

> ENFSI wishes to promote consistent and reliable scientific information through the whole forensic process from the scene of crime to court. It recognizes that it is of vital importance to interpret potential forensic evidence in the context of a case, using all available information; reporting results of analysis alone may be misleading.
>
> The ENFSI Board plans to undertake actions to agree a standard for the interpretation of forensic evidence and to provide the necessary support for implementing this standard within the membership (European Network of Forensic Science Institutes, 2010).

This call for a coordinated position regarding issues of forensic interpretation across Europe was preceded by initiatives such as the UK-based project for the implementation of a harmonised interpretation standard under the auspices of the Association of Forensic Science Providers (AFSP), published in 2009 (Association of Forensic Science Providers, 2009). This document encodes general principles for designing cost-effective examinations and interpretation of results of forensic examinations as published since the late 1990s—the framework also known as Case Assessment and Interpretation (CAI) (Cook et al., 1998a, 1998b; Jackson and Jones, 2009). More recently, this fundamental understanding has been restated and thematically organised in a series of practitioner guides, issued by the Royal Statistical Society's Statistics and the Law Working Group (Aitken et al., 2010). The practical implementation of such generic principles encounters several hurdles, however, among which is the commercialisation of forensic services. For example, in his comment on the situation in England and Wales, Jackson (2014: 16) notes that budgetary considerations may deter a request for an advanced level of interpretation, even though such a level may be crucial to ensure that the parties at trial will be properly assisted in their task of weighing the scientific findings.

The instances of organisational commitment to forensic interpretation discussed above are only selected examples, however, and it is important to emphasise that they are not entirely supported in the scientific community, as is shown by various opposing opinions raised in scientific literature and academic debate. Discussions about scientific evidence evaluation have a long history in print (Taroni et al., 1998), but for the ease of argument, consider a few instances from the past two decades.

In his seminal paper published in the early 1990s, Stoney set out the challenge by asking what made us ever think we could individualise using statistics (Stoney, 1991), while Evett warned against purely opinion-based positions when he wrote: 'Don't fall for the ''when you've been doing the job as long as I have'' line' (Evett, 1996). About a decade later, Saks and Koehler (2005) critisised the fact that, besides some exceptions in the forensic DNA analysis, forensic science as a whole still largely relies on error-prone heuristics. They expressed the hope for a 'paradigm shift in forensic identification science', yet a decade later it remains questionable if or when this shift is going to come. This is illustrated by the fact that it remains common to read statements such as '[t]his uniqueness is obvious to anyone who spends all day looking at medical X-rays' (Quatrehomme et al., 2014) in peer-reviewed journals, which is in sharp contradiction to Evett's quote above, as well as the results of comprehensive enquiries about the notion of individualisation (Cole, 2009, 2014).

In combination, the above demonstrates the existence of disparate practice and opinions across the forensic science community regarding aspects of interpretation. This is problematic as it conveys a picture of vagueness among recipients of expert information and raises doubts about the extent to which forensic science services can be trusted. At a time when criminal justice authorities of EU countries engage in dialogue and action to build a common area of justice and to help increase the mutual trust in judicial systems across Europe, the agenda of the forensic science community should place interpretation in a prominent position. ENFSI has committed itself to this topic as part of its Monopoly Programmes (MP), which regroup thematic clusters of projects funded through the 'Prevention of and Fight against Crime' program of European Commission. Forensic interpretation was a main theme of the 2010 MP Programme 'Strengthening the Evaluation of Forensic Results across Europe (STEOFRAE)'. In particular, one of the projects that came under this umbrella, entitled 'The development and implementation of an ENFSI standard for reporting evaluative forensic evidence' (or M1 for short because it is the first in a

series of seven projects), focused on the encoding of cross-institutional understanding about generic principles of forensic interpretation, the specification of a roadmap for implementation and the provision of educational support on this topic for practitioners across Europe.

In this paper, we will first outline the key elements of the M1 project, followed by a review of some commonly raised issues related to selected aspects of the endorsed principles of interpretation. This discussion is based on the authors' experience through peer consultations with forensic science practitioners. After a substantial effort towards the establishment of that standard that became a guideline, we felt the need to take a step back to try to identify the key learning points of the initiative for us. We invite the reader to consider this a personal exercise, and all opinions expressed are ours only and do not involve the other members of the project's core group (see Appendix), working groups and practitioners, who all contributed to the discussion. They all fuelled our thinking for this paper and we thank them. As a main standpoint, we will seek to justify the view that addressing topics in forensic interpretation *now*, despite implementation challenges, is of vital importance because forensic science continues to be challenged by proactive participants in the legal process and a general context that tends to become more demanding and less forgiving (for example, the public and the media). Further conclusions will be presented in the final section.

The ENFSI M1 project

Background and history

The ENFSI 2010 MP project 'The development and implementation of an ENFSI standard for reporting evaluative forensic evidence' (M1) arose at a time when the need to implement good practice for forensic interpretation and reporting was recognised, but difficult to address because of very diverse practices both within and across ENFSI laboratories. Although ENFSI working groups in particular areas of forensic practice have worked individually on domain-specific guidance documents and recommendations, there was no generally agreed statement from the community as a whole on generic principles to which forensic evaluation and reporting ought to conform. The M1 project focused on filling this gap, which is considered essential as it can provide a framework within which community-wide changes can be enabled with coherence. The major project aim was thus the development of an agreed reference document regarding fundamental principles of forensic interpretation independent of the forensic area of application. The project core group, led by Dr Sheila Willis (Forensic Science Ireland, Dublin, Ireland), was composed of experienced caseworkers and recognised academics from ENFSI institutions across Europe. The works of the core group (36 months, starting in 2011) were regulated by a so-called 'Product development process' document, which specified key development stages such as consultation in relevant expert communities, the collection of feedback, the revision of draft versions of the reference document and possibilities to appeal against core group replies on specific comments or requests for changes.[1]

Brief overview of the core principles of the M1 document

The overarching aim of the M1 document is to ensure that forensic practice will focus on those activities, and only those, that help resolve the key issues in the case, save costs for the submitting party, make efficient use of resources and conform to the criteria balance, logic, robustness and transparency. To achieve this aim, the forensic practitioner must strive to understand the needs of the submitting party and the key issues in the case. This requires relevant information, which should be made available or requested if needed. It is worthy of emphasis, however, that not all but only *relevant* case information

1. The final document is available at the following links: http://www.unil.ch/esc/home/menuinst/les-unites/police scientifique/collaborations.html and www.enfsi.eu/news/enfsi-guideline-evaluative-reporting-forensic-science (accessed 24 October 2016).

is required, that is, information that has a bearing on the sound assessment of the strength of the scientific findings. In both theory and practice, information became to be recognised as one of the fundamental ingredients that condition rational reasoning schemes, so that it is considered mandatory to state the available information explicitly. Further, reports are required to mention that in the event of changes in the conditioning information, the adequacy of competing propositions, assessments and conclusions need to be reviewed.

However, once the needs of the submitting party are elicited, the physical examination process may *not* start immediately. It is necessary first to assess the extent to which any examinations to be carried out could lead to results that help discriminate between the competing propositions of interest. This process of reasoning, known in literature on CAI as pre-assessment (Cook et al., 1998a), places evaluative considerations at the beginning of the case. This may be of particular interest for cases in which examinations may be cost-intensive and time-consuming (certain types of transfer traces, for example) or require an evaluation at a higher level in the hierarchy of propositions (activity level, for example). In fact, if pre-assessment shows that, whatever the outcome, the strength of the evidence could not be assigned (for instance because specialised knowledge on factors such as background and phenomena of transfer and persistence is unavailable), the instructing party may take the informed decision not to commission any examinations, but rather to review its priorities and allocate its resources to other aspects of the case.

The M1 document promotes an overall thinking process that directs the evaluator's attention to the findings *given* at least a pair of competing propositions. To ensure an assessment of the scientific findings that is tailored to the case, propositions must derive from relevant and agreed case information. In particular, the alternative proposition should reasonably reflect the position of the defence by taking into account, for example, statements made by the person of interest, as illustrated in the case example outlined earlier.

Given conditioning information and clearly defined propositions as outlined above, the core part of the evaluative process can be carried out. It amounts to a consideration of the scientific results *given* the various competing propositions and relevant conditioning information (and not the reverse), expressed in terms of conditional probabilities. That is, it is considered on the one hand how probable the findings are given both the first proposition and the conditioning information, and on the other hand how probable the same findings are given the alternative position and the conditioning information. Comparing these two terms provides an expression of the strength of the scientific results, more commonly known as the likelihood ratio (Robertson and Vignaux, 1995; Aitken and Taroni, 2004). To enable an audit of the way in which the likelihood ratio was assigned, the case file must document the component conditional probabilities and any data that was used in the evaluative process.

The likelihood ratio thus represents an essential core part of the evaluative process, and forms the basis for the conclusion of the evaluative report. It contributes to a balanced approach because the findings are considered from at least two competing points of view (i.e. propositions). The likelihood ratio can be conveyed either by its value or complemented by a verbal equivalent, that is, a statement about which, if any, of the competing propositions is supported by findings. The reported strength of support informs the recipients of expert information how they ought to revise their beliefs about the competing propositions, whatever their beliefs were prior to consideration of the scientific findings.

Document structure

The key elements of the guideline document as outlined in the previous section are encoded in a compact form with three sections—'Scope', 'Evaluative reporting' and 'Standard framework'—covering a total of less than two dozen paragraphs. The remaining parts of the guideline document, that is, about three quarters, cover a section with four 'guidance notes' and a substantial glossary. The four guidance notes cover the topics 'Reporting requirements' (Guidance Note 1), 'Propositions' (Guidance Note 2), 'Data and expert knowledge used to assess the strength of the findings and assignment of likelihood ratios'

(Guidance Note 3) and 'Meaning of the likelihood ratio in an evaluative report' (Guidance Note 4). Despite the extensive explanations provided in these Guidance Notes, and the definitions given in the glossary, the authors' experience through consultations with forensic practitioners shows that there is a considerable gap between, on the one hand, the intentions of a harmonised view on principles of forensic interpretation and, on the other hand, the way in which efforts to progress towards such common understanding through the M1 document is perceived in the community. In the next section we will review and discuss several recurrently raised concerns.

Gaps between intentions and perceptions of key concepts

Evaluative reporting. The scope of evaluative reporting, as it is understood in the M1 guideline, is strictly related to a likelihood ratio-based assessment of probative value, in the context of alleged circumstances, and the statement of such an assessment in a communication to a recipient of expert information. Defining evaluative reporting in terms of the likelihood ratio has the advantage of providing explicit criteria to judge the nature of a report. The likelihood ratio format achieves this through the following three key elements: conditioning information, at least a pair of propositions, and focus on the probability of the findings *given* each of the competing propositions (Evett and Weir, 1998). Among the most common situations where this reporting structure arises are those where comparative examinations have been carried out between material of unknown source and reference material from one or several potential source(s). Here, the fact that there is a least one potential source provides obvious grounds for formulating a first proposition regarding the origin of the question material, or the activities that led to the production of the recovered trace material. To assess the way in which the results of the comparison should impact on the mandating party's view of this first proposition, general rules of logic require the specification of at least one alternative proposition.

This scope for *evaluative* reporting—strictly defined in terms of a likelihood ratio—has been questioned on the ground that, in many other settings where forensic examinations are carried out, expert conclusions are provided based on what is also called 'evaluation'. For example, a footwear mark specialist may examine a mark from a crime scene to provide indications about the model and make of the shoe that left the mark. Clearly, there is no potential source at this juncture and hence there are no formal propositions among which the scientist could help to discriminate. Instead, the examiner would typically offer an expert opinion on potential explanations. In the literature on CAI, such expert activity has been recognised as an *investigative mode* of functioning (Jackson et al., 2006). It usually arises at an early state of investigative proceedings, involving no potential source (such as an object or a person of interest), and there is no risk that the expert's expression of an opinion on a proposition will conflict with the reasoning processes of other participants in the legal process. Such reporting is thus referred to as investigative reporting as it exhibits fundamental differences compared to the evaluative setting defined above.

Further instances of forensic reporting that are not evaluative in the sense of the M1 guideline, described as technical or factual reporting, involve cases where examiners provide an essentially descriptive account of findings. This may take the form of statements about observations, such as the nature of particular physical matter, but also the assignment of objects to classes (i.e. classification). For example, forensic examiners may describe the morphological features of a fibre or analyse powder of unknown substance to characterise its qualitative and quantitative composition. Such technical reporting may contain elements also sometimes referred to as '(statistical) evaluation', typically descriptive statistics for summarising series of quantitative measurements, but this does not correspond to the notion of evaluation under a pair of competing propositions in terms of a likelihood ratio. For example, the reporting that an unknown submitted substance is heroin with a concentration of 23% ($\pm$ 5%) is a typical case of that type of reporting.

Distinguishing investigative and technical reporting from evaluative reporting is not meant to introduce a breach in the value of different reporting formats, but to acknowledge differences in the purpose,

nature of the conclusions and the argumentative grounds to reach conclusions. In fact, evaluative reporting is often preceded by or contains elements of technical reporting because these provide the descriptive account of the findings, which is a necessary preliminary to evaluative reporting.

Propositions. The concept of propositions is commonly objected to from several perspectives. First and foremost, it is argued that 'dealing with' or 'addressing' propositions is out of the question for forensic experts as this may impede on their role, compared to that of the competing parties or the legal decision maker at trial. This would indeed be the case where experts express an opinion on particular propositions, but this is not what evaluative reporting entails. The document is not giving any invitation to address what is referred to as the ultimate issue. As previously noted, evaluative reporting uses propositions only as a conditioning. That is, the focus is on the probability of the findings *given* particular propositions and conditioning information. Stated otherwise, forensic experts do *not* address propositions as such, but rather the findings under the assumption of propositions of interest. Responses and interactions in the course of the project highlighted that there is a lack of understanding about the difference between addressing the findings and addressing the propositions, and this lack of understanding may be the source of some of the resistance.

Related to this is the objection that, in many cases, formal propositions are not available to the forensic expert, or may continuously change as the case moves on in the process. Moreover, on a procedural account, the defence may be under no obligation to make a statement. In practice, the defence might also choose to reveal its position only once the case is at trial. Forensic experts thus consider the absence of propositions, in particular the absence of an alternative proposition, as a major hindrance for evaluative reporting. However, this is an operational constraint and interferes with the establishment of evaluative reports, but it cannot bend the principles of evaluative reporting. To avoid an impasse, the M1 guideline proposes the option of adopting an alternative proposition that most reasonably reflects the party's position, in view of the available circumstantial information. Although this might seem as an invasive approach, it is the only option that can lead to an evaluative report. It is required, however, that reports specify that any change to the propositions, in particular any new propositions proposed by the parties or mandating authority, may impact on the assessment of the strength of the forensic findings. The extent to which this is the case will need to be evaluated in the case at hand and may lead to the provision of a revised report. Alternative ways to deal with the absence of propositions are the exploration of explanations in an investigative report and the restriction to technical reporting.

The concept of propositions is also viewed cautiously because of its apparent association with the idea of 'explaining findings away'—that is, working under a propositional framework is considered as an invitation to come up with any alternative explanation for the findings. This paralyses the evaluation of probative strength when the probability of the findings given an explanation is the same, or about the same, as the probability of the findings given the first proposition, so that the resulting likelihood ratio will be about one. Although such an assessment considers two opposing sides, and thus is balanced, it will be uninformative and of no assistance to the court. This complication can be settled by distinguishing explanations from propositions, as emphasised in existing literature on CAI (Evett et al., 2000; Jackson et al., 2013). Unlike propositions that can be justified on a background of relevant case information, explanations may be a statement of the obvious, speculative or fanciful, without offering a reasonable conditioning for probability assignment. For these reasons, the M1 guideline supports evaluation based on propositions. A mere listing of alternatives that have the potential to explain findings without consideration of whether such explanations find a reasonable justification with respect to the circumstances of the case, as is widely done in areas such as gunshot residue (GSR) analysis,[2] is

2. GSR analysts commonly report on how to explain (i) the presence of particles even though a person of interest did not discharge a firearm, or (ii) the absence of particles even though a person of interest discharged a firearm. Thus, whatever the quantity of recovered particles, an uninformative 'could be' or 'consistent with' is reported. See, e.g., Gallidabino et al. (2015).

not endorsed. Further examples for explanations are general statements such as: 'The injuries are consistent with having been caused by the weapon found on the crime scene', 'The defendant cannot be excluded as a source of the partial DNA profile seen in the mixture of DNA on the swabs' or 'The shoe mark found at the point of entry could have been made by the suspect's shoe.'

An important concept regarding propositions is that they can be classified into broad categories, so-called hierarchical levels (Cook et al., 1998a), such as 'crime level' (propositions that refer to the commission of a criminal offence), 'activity level' (propositions about a human activity or a happening), and 'source level' (propositions about the source of physical matter). While the ENFSI guideline is supporting these distinctions, it also seeks to counter two common misconceptions: first, many forensic practitioners think that source level propositions suffice, and secondly, many magistrates believe that moving up to activity level is usurping the duty of the court. Regarding the first point, the guideline emphasises that source level propositions are *not* sufficient and can be misleading. Thus, activity-level propositions *should* be considered, when (i) factors such as transfer mechanisms, persistence and background levels of the trace material affect the understanding of scientific findings relative to the alleged activities, and (ii) when the appropriate assessment of such factors requires expert knowledge. This observation is of primary importance for trace materials such as fibres, glass, gunshot residues and small quantities of DNA, drugs or explosives. Regarding the second point, there is no breach with respect to the duty of the court as long as the scientist focuses on the findings given the propositions, and not the reverse (see also 'The probability of the findings given the proposition and concerns regarding data' below).

Circumstantial information. The common objections against propositions mentioned in the previous section, in particular availability and continuing changes as a case moves on, are similarly levelled against the requirement for relevant circumstantial information. As for propositions, however, these challenges represent issues of practical application that do not adversely affect the crucial role of information as an element of a logical approach. A more unfortunate reproach to information, in its more general sense, is the risk of contextual bias of experts' expectations and their opinions, a topic that has caused controversial discussion (Buckleton et al., 2014; Champod, 2014b; Risinger et al., 2014). To avoid a digression to this topic, the emphasis is only on *relevant* information as emphasised in the previous section, that will help the examiner to understand the needs, choose the appropriate techniques and assign the probability of the findings given the competing propositions (Jackson et al., 2013). Clearly, for example, the examiner does not need to know if the person of interest has previously been convicted or has been 'identified' by an eyewitness. However, it may be relevant for the examiner to learn about the background of the person of interest in cases involving, for example, certain types of transferable particles, such as glass or GSR. The reason for this is that, under the assumption that the person of interest denies any involvement in the case, the assessment of background quantities of target particles should take into account the fact that certain professional activities (demolition works, for example) lead to increased quantities of target material.[3] Similarly, information about timing of events is crucial to help ensure that results of toxicological analyses are understood appropriately in the context of the case as a whole (Willis, 2014). Yet another example is selected trace material, such as a small cut of clothing bearing staining, without any information about the exact location of the staining (on the clothing), its abundance and distribution. In such a case, the examiner might at best help inform about the source of the recovered material, but the real need of the recipient of expert information may lay beyond this level of propositions. Indeed, the fact-finder might be interested in so-called activity-level propositions to which the examiner's specialised knowledge about phenomena such as transfer, based on information relating to

3. In the context of glass particles, for example, this can lead to more moderate assignments of evidential strength, and hence more favourable assessments for the defendant. See, e.g., Curran et al. (2000). To ensure a fair weighing of the findings, it is thus desirable that examiners are allowed access to such relevant information.

distributional patterns, may provide crucial insight. While the need for relevant information may appear obvious in such cases, it turns out that in some jurisdictions budgetary considerations, commercialisation of forensic services and the lack of understanding of the need tend to limit the number and quantity of trace items submitted to laboratories, thus hindering vital discussion from being raised in case-based assessment (Jackson, 2014).

As noted in the above example, relevant case information not only helps direct the attention to levels of propositions that more closely approach the submitting party's needs. Appropriately framed information can also help avoid analyses being carried out or reported that will add little value in discriminating between competing scenarios. In an assault case, for instance, there might be little interest in corresponding DNA profiles if the person of interest admits social contact with the victim prior to the alleged incident, or claims to have helped the victim. It is not uncommon, however, to encounter experts who would insist on reporting about corresponding DNA profiles in such cases by means of naked rarity figures in the order of, say, one in the billions, and accept the danger that such an impression of strength of support for source-level propositions might be inappropriately carried over to propositions at a higher level (i.e., activity or crime level). Justifications for such practice are not always quite clear, but stem from the archaic view that corresponding DNA profiles are to be accompanied by conditional genotype probabilities, and that reporting of such figures replies to expectations of the judiciary.

It is also worth noting that a number of forensic examiners regard the topic of relevant circumstantial information sceptically because of a misconception of their role and relationship with information. That is, examiners commonly mention that they 'don't *know* if the information is true' and that 'it might turn out to be inaccurate in various ways'.[4] The misconception here is that experts are not expressing an opinion about the circumstantial information, nor are they held responsible for the accuracy of the information. The information is the one that has been presented to them at a given instance in the procedure, and should any of this information change, then the evaluation will need to be reviewed. Such a revision is not a corrective action in response to an 'error', but a natural reconsideration in the light of a new situation.

The above observations point out that there is room for improvement of the understanding between both forensic experts and recipients of expert information that pertinent case information has the potential to unlock the capacity of forensic findings to add value to the evaluative process. To achieve this, it is necessary to overcome the widespread perception that sees forensic science as a practice that can lead to results that 'speak for themselves' and carry an intrinsic value. This is not the case, except for instances where forensic accounts are exclusively descriptive, and hence not evaluative. Instead, it is generally understood that the elicitation and assignment of probative value requires the specification of an inferential target, formulated in terms of propositions, conditioned on a background of information (Robertson and Vignaux, 1998). Further, it should be emphasised that, usually, someone *is* working on some version of the circumstances. This may be the submitting party or the scientist who decided which items to prioritise. Therefore, it is more transparent to make the information on which these decisions are based explicit.

The probability of the findings given the proposition and concerns regarding data. In his presentation 'Clarity or confusion? Making expert opinions make sense', held in 2009 on the occasion of the celebration of the 100th anniversary of the School of Criminal Justice of the University of Lausanne, Ian W Evett noted that '. . . the single most important advance in thinking about interpretation of evidence is the notion that it is the scientist's place to talk about the probability of the evidence given the proposition, *never* the probability of the proposition.' In forensic and legal literature, the probability of the findings *given* propositions has been known for a long time as one of the principles of interpretation (Evett and Weir,

4. A similar objection is sometimes encountered with respect to propositions when examiners argue that they don't *know* if a particular proposition is accurate. See Evett et al. (2000).

1998; Finkelstein and Fairley, 1970), but experience shows that it is still widely misconceived and treated with scepticism in various ways.

One major objection is not actually levelled against the conditioning on propositions, but against probability in general. The objection consists of, first, equating the term 'probability' with 'statistics', secondly, arguing that statistics cannot be done without data, and thirdly, concluding that the absence of appropriate forensic data prevents probability from being used in forensic practice. There are several closely intertwined aspects in this chain of argument that are worthy of consideration. First and foremost, probability—understood in a general sense—deals with the reasonable reasoning in the face of uncertainty. Statistics has, indeed, been described as the study of uncertainty (Lindley, 2000), but this is not the feature highlighted in the objection mentioned above. The objection appears to focus statistics exclusively on data analysis, which, to some extent, involves descriptive and exploratory techniques that obviously depend on available data. Although such analyses make reference to uncertainty, statistics also deal with uncertainty away from data (Lindley, 2000). It arises, most typically, when using data to draw inferences about competing propositions. Therefore, scarce or missing data, a characterising feature of forensic science, do not prevent the appearance of uncertainty in relation to topics beyond data and the possibility of their study through probability as an integral part of statistics.

A variation of the above objection maintains that, without data, one cannot 'give' a probability.[5] The suggestion underlying this strain of concern appears to be that probability necessarily requires data, in particular so-called frequency data, and thinking in terms of near-identical repetitions. This is a flawed and extremely widespread perception that can be traced back to the frequency definition of probability.[6]

Practically, this represents a problem because most educational programmes throughout the world and across disciplines are essentially based on the frequentist approach to statistics. Its relevance for reasoning in forensic science and the law is compromised on both theoretical and practical grounds. On a theoretical account, there is a terminological issue in that the notions of frequency and relative frequency relate to data,[7] whereas probability is properly reserved for an individual's measure of uncertainty (for example, about a proposition of interest). Clearly, there is a relation between data and probability, but two are not the same. The former, data, can inform the latter, probability, but the absence of the former does not mean that no probability could be assigned. Indeed, most if not all real-world events of the present, the past and the future which entail aspects about which a reasoning individual is uncertain cannot be meaningfully reconstructed in frequentist terms. This represents an applicability problem for the frequency-based approach and is one reason why it is often avoided in favour of alternative views, such as the belief-type interpretation of probability (Lindley, 1991). The latter view of probability, also called the personalist or subjectivist approach, does not encounter conceptual limitations with respect to singular, non-repeatable events. This makes it particularly attractive from an operational point of view (Lad, 1996). Most importantly, probability assignment in the subjectivist perspective is more flexible regarding the data at one's disposal. Hence, limitations in the availability of data cannot serve as an argument against probability, only against a particular definition of probability. A possible reason for difficulties with this understanding might be that the primary education of many forensic practitioners is in the analytical science fields, which may direct their attention to instrumentation rather than probability.

The M1 guideline document addresses the above challenges by allowing probability to be understood from a subjectivist perspective. This enables forensic experts to base probability assignments on their full body of expert knowledge. This acknowledges the fact that probability involves more than data (in the sense used before) alone. For example, forensic experts may have specialised knowledge about the

5. For a discussion of this objection on behalf of fingerprint examiners, see the section on 'Categorical statements'.
6. For a critique of frequentist ideas in the context of forensic DNA profiling results see, e.g., Brenner (2010). Discussions on general foundational aspects can be found in, e.g., Lindley (1985, 2014).
7. Frequency refers to the number (i.e. counts) of observations of a particular kind, whereas the relative frequency is the number of observations of a particular kind divided by the total number of observations made.

manufacturing process of particular items, such as shoe soles and tools, that is relevant for the assessment of the occurrence and persistence of observable features (such as marks and traits). In turn, DNA analysts may invoke established genetic theory when assessing the rarity of selected genetic features (Brenner, 2010; Buckleton et al., 2005; Evett and Weir, 1998), or specialised knowledge about phenomena of DNA transfer in various situations. The availability of such bodies of knowledge for auditing and disclosure will ensure that subjective probability assignments are not vulnerable to the common critiques that label them as guesswork, arbitrary or speculation (Champod, 2014b).

Understanding probability in a subjectivist perspective is both an opportunity and a challenge. The opportunity is that relevant expert knowledge that is not necessarily encoded in case-specific *hard data* can be beneficially brought into the process. As an example, consider the case *Barry George* v *R*,[8] where the finding consisted of one particle (GSR, a.k.a. FDR in the UK) in the pocket of a coat, one year after the shooting of the victim. There is no reasonably feasible way to address the probability of this finding in terms of 'frequency data' or to design experiments to simulate cases with time lapses over a year. But still, there is expert knowledge in this particular area of forensic practice regarding the nature, formation, quantity, transfer and persistence of particles as a result of the discharge of a firearm. The challenge is to use such understanding as a basis for the expert's assessment of the probability of the findings given the competing propositions, and the subsequent comparison of the assigned conditional probabilities. That is, the demanding task for experts is to elicit their specialised knowledge and construct a scientific argument that will allow them to convey their assessment, as illustrated, for example, in the following transcriptions:

> In our opinion, it would be just as likely that a single particle of discharge residue would have been recovered from the pocket of Mr (. . .)'s coat whether or not he was the person who shot Ms (. . .) nearly a year previously.[9]

> In our opinion the probability of finding a single particle of discharge residue in Mr (. . .)'s coat pocket would have been the same, regardless of which of the above propositions was true. The FDR evidence is thus inconclusive. In our opinion it provides no assistance to anyone asked to judge which proposition is true.[10]

Restricting evaluative reporting only to cases where hard data are available would imply that the scientist retreats to a technical reporting[11] that is limited to a statement of the findings in a descriptive way, with no indication as to how the recipient of experts expert information ought to use the findings to revise their opinions regarding the competing propositions, as demonstrated in the quotes above. This would be detrimental to a variety of forensic areas of practice (comparative handwriting examinations, for example) that similarly rely on the comparative assessment of subjective probabilities given competing propositions and lead to the possibility that the findings will be misinterpreted in the absence of some aspect of expert knowledge.

Categorical statements. Probability assignment for findings given observations, discussed in the previous section, receives a particular challenge from adherents of the traditional identification/individualisation disciplines, such as fingerprint examination. They argue, first, that no 'validated probabilistic model' currently exists that could account for the entire scope of forms in which ridge skin impressions come along, and that could be used in a reliable way to make probability statements. Secondly, given the inability to assign probabilities, the argument maintains that the accustomed conclusions 'exclusion', 'identification' and 'no interpretation/inconclusive' remain appropriate and must continue to be used.

8. *Barry George* v *R* [2007] EWCA Crim 2722 (15 November 2007).
9. Ibid. at [22].
10. Ibid. at [23].
11. See also a mere listing of potential explanations would not be a viable alternative reporting format either [2].

This critique regarding non-existent probabilistic models is unfounded in the current state of research (Neumann et al., 2012) and its conclusion—the requirement to adhere to categorical conclusions—does not follow from it *a fortiori* if probabilistic models were in fact non-existent. That is, if one could *not* assign probabilities, and hence assign no likelihood ratio for the result of a ridge skin comparison, then this means that, first, no evaluation of strength of support could be reached, and secondly, any stronger conclusion beyond strength of support—most notably categorical conclusions (such as individualisation)—*cannot* logically be sustained either.

Although this might not prevent practitioners from adhering to a reporting practice with categorical conclusions, the fundamental problem known as the 'leap of faith' (Stoney, 1991) persists. This is unsatisfactory because it is an unscientific position and hence cannot become part of a guideline document. More generally, whatever one's position regarding categorical conclusions, the concept of evaluation as understood in the M1 document[12] remains an inevitable logical preliminary to conclusions that reach beyond expressions of strength of support (Biedermann et al., 2008).

For the above reasons, the M1 document considers categorical conclusions regarding common source (for example, in the examination of marks on elements of ammunition, handwriting, fingerprints, footwear marks, etc.) to sit outside its scope. The guideline document merely formalises the fundamental considerations that any evaluator must undertake if the purpose is to assist in the process of inference of identity of source, whoever is going to make the ultimate decision. Seriously committed examiners already endorse this perspective by considering their findings (features in a ridge skin fingermark, for instance) not only from the perspective of the mark coming from a given person of interest, but also from at least one alternative, such as the mark coming from some unknown person. It is worth reiterating at this juncture that it is this framework of reasoning that represents the core of the approach, and not the availability of data, as argued in the previous section, or a statistical model. There is nothing in the M1 guideline that prevents an examiner from reasoning about the probability of a given general pattern or the probability of encountering a given configuration of minutiæ in a given ridge skin area without a computerised model. There is a substantial body of knowledge and accepted theory (on morphogenesis, for example) in the field to support such probability statements. Without such expert knowledge, both explicit and implicit, the currently reported conclusions regarding identity of source could not be made either.

Aversions against the notion of standard. Originally, the ENFSI M1-project included the term 'standard' in its title ('The development and implementation of an ENFSI standard for reporting evaluative forensic evidence'), and so does the ENFSI statement regarding interpretation (European Network of Forensic Science Institutes, 2010). Consideration of the M1 document as a standard has been, however, one of the most recurrent source of objections, although this seems contrary to the widespread consensus in the field of forensic science and the law that sound principles of forensic interpretation are an important point in the agenda of works towards common understanding and a smooth functioning of judicial systems across Europe. Many commentators even made their support for the final document dependent on a change to a weaker and less enforcing status, such as a guideline or recommendation. Aversions against the notion of standard have also been expressed forcefully in literature by classifying communications on ongoing standardisation works as 'lobbyism' and by favouring, instead, pluralistic perspectives (see, for example, Simmross, 2014). Representatives of some expert fields, such as dactyloscopy, dismiss the proposed evaluative framework outlined in 'Brief overview of the core principles of the M1 document' and 'Evaluative reporting' above as being incompatible with their existing reporting conventions. The extent of this skepticism at different fronts appears at odds with the fact that forensic science has no difficulty in adopting standards on a variety of operational topics. Typical examples include the ASTM Forensic Science Standards[13] on methodological aspects of examinations in various

12. See the section on 'Evaluative reporting'.
13. Available at www.astm.org/Standards/forensic-science-standards.html (accessed 24 October 2016).

forensic domains. While these are rather technical considerations regarding laboratory practice, standards have also been issued on evaluative aspects in some forensic areas regarding, for example, inference of common source.[14] Further examples in this context are the forensic standards issued by the Australia Forensic Analysis Committee, in particular documents AS 5388.3 on interpretation and AS 5388.4 on reporting.[15]

This apparent inequality in the perception of technical and inferential standards might stem from the fact that standards regarding technical matters are far more easily adopted, as they require little more than adequate laboratory equipment and trained personnel capable to operate relevant machinery according to predefined standard operational procedures. Evaluative standards, in turn, are crucially depending on the intellectual capacities of the thinking scientist. This reinforces the view that forensic science is a discipline that defines itself through the support it provides to other participants in the legal process, in particular in matters regarding the inferences that may be drawn on the basis of scientific findings. This is a view of the nature of forensic science that Evett has expressed as follows:

> . . . I will settle for a simple premise: forensic science is a state of mind, I mean that whether a particular individual is behaving, at a given juncture, as a scientist can be determined by the mental process underlying his/her actions and words. (Evett, 2010: 121)

Interestingly, none of the above scepticisms are directed against the logic of the proposed standard framework of reasoning. Instead, the disapproval seems to rely merely on an argument to the people, that is, considering the proposed evaluative principles to be inappropriate because they run contrary to the majority of current perceptions and because they would be applied to a community that is not prepared or not proficient and knowledgeable enough to follow the approach. Since the prevalent resistance seemed to be focused more on the title of the document, rather than its content, and broadening the support across the community is desirable, the development of the M1 document was pursued by labelling it a guideline, rather than a standard. This concession was deemed acceptable not least because the core content of the document is rigorous, based on sound scientific principles and unique in its final form throughout Europe. Thus, it has the potential to become considered, *de facto*, as a standard.[16]

Conclusions: envisioning benefits and changes

The M1 reporting scheme is based on principles that have previously been operationalised in terms of, for example, the CAI model (Cook et al.,1998a), reviewed and refined over years and encoded as a standard by forensic science providers (Association of Forensic Science Providers, 2009). This should make the scheme both practical and achievable. Notwithstanding this, the field's readiness to implement the proposed framework varies across institutions and jurisdictions, depending on limitations in practical capacities that may represent serious hurdles. This situation is further complicated by the fact that many forensic agencies face budgetary constraints that limit their possibilities to train key personnel on topics in forensic interpretation. The core part of the M1 document cannot cover this gap because it is not intended as a training manual. It intends to ensure comprehensiveness through a substantial glossary and to provide guidance and assistance for implementation through supporting documents such as a proposal for a roadmap, an audit template and examples of case reports formulated according to the guideline. Other budgetary constraints are subtler. To the best of the authors' knowledge, there are no recognised

14. See, e.g., Bundeskriminalamt (ed.) (2010) *Standard of Dactyloscopic Proof*. (Referat KI 32) [translation of: *Standard des daktyloskopischen Identitätsnachweises*].
15. These documents—not endorsed by the authors—are mentioned solely for the purpose of their general thematic relationship with the topic discussed in this paper.
16. As an aside, it is worth mentioning that labelling the M1 interpretation principles as 'European' is a slight overstatement because the underlying understanding has previously been supported by scholars from other parts in the world, as far as New Zealand. See, e.g., Robertson and Vignaux (1995).

performance indicators for forensic science other than case numbers or outputs. Therefore there is a trend to increase the case numbers to 'grow' the budget. The authors' own experience is that it may be difficult to implement the guideline document and one of the constraints is the reduction in outputs that results.

The development of a consensus document regarding generic principles of forensic interpretation not only represents an ambitious aim, but also comes at a cost. This makes it relevant to enquire about the potential of the initiative to add value and to lead to net benefit. The main reply to this is to consider the key achievement of a consensus document that will consist of establishing a common language to define criteria for a rational assessment of scientific findings. Such a document will set out specifications designed to ensure that the expert's reasoning and subsequent reporting is scientifically sound, balanced, transparent and can be subjected to an audit. The benefit of such a framework is twofold. On the one hand, defendants will be able to expect that scientific results regarding their case *will* be assessed according to agreed principles and not merely by an approach that leaves the choice of the evaluative scheme to the discretion of the expert (Biedermann et al., 2015). This will ensure equality of treatment and help avoid vital considerations in the assessment of scientific evidence being kept from defence. On the other hand, forensic experts will benefit from procedural consistency across forensic disciplines. This will enable them to concentrate on the core questions of interpretation, that is, the assignment of likelihood ratios and explaining the grounds for their assessment in the case at hand, rather than deploying efforts to repeatedly justify the first principles of forensic interpretation. By taking up this topic *now*, and envisioning a change, forensic specialists have an opportunity to take a proactive role in the shaping of a core part of their expert activity. Providing a clear definition of the scope, nature and meaning of evaluative reporting will help shape expectations among other participants in the legal process. Thus, at a time when recipients of expert information tend to become more demanding and less forgiving, increased mutual understanding regarding generic principles of consistent reporting followed by experts should also contribute to reducing the scope of the attacks to which expert reporting is commonly exposed.

Declaration of Conflicting Interests

The author(s) declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

Funding

The author(s) disclosed receipt of the following financial support for the research, authorship, and/or publication of this article: The works reported in this paper have been financially supported by the Prevention of and Fight against Crime, Programme of the European Union European Commission—Directorate—General Justice, Freedom and Security, EU ISEC 2010, Agreement Number: HOME/2010/ISEC/MO/4000001759. Alex Biedermann was supported by a Grant of the Fondation pour l'Université de Lausanne and the Swiss National Science Foundation (Grant No. BSSGI0_155809).

References

Aitken C and Taroni F (2004) *Statistics and the Evaluation of Evidence for Forensic Scientists*. Chichester: John Wiley & Sons.

Aitken C, Roberts P and Jackson G (2010) *Fundamentals of Probability and Statistical Evidence in Criminal Proceedings (Practitioner Guide No. 1): Guidance for Judges, Lawyers, Forensic Scientists and Expert Witnesses*. Royal Statistical Society's Working Group on Statistics and the Law. Available at: www.rss.org.uk/Images/PDF/influencing-change/rss-fundamentals-probability-statistical-evidence.pdf (accessed 24 October 2016).

Association of Forensic Science Providers (2009) Standards for the formulation of evaluative forensic science expert opinion. *Science and Justice* 49(3): 161–164.

Biedermann A, Bozza S and Taroni F (2008) Decision theoretic properties of forensic identification: underlying logic and argumentative implications. *Forensic Science International* 177(2–3): 120–132.

Biedermann A, Vuille J, Taroni F and Champod V (2015) The need for reporting standards in forensic science. *Law, Probability and Risk* 14(2): 169–173.

Brenner C (2010) Fundamental problem of forensic mathematics: The evidential value of a rare haplotype. *Forensic Science International: Genetics* 4(5): 281–291.

Buckleton J, Triggs C and Walsh S (2005) *Forensic DNA Evidence Interpretation*. Boca Raton, FL: CRC Press.

Buckleton J, Evett I and Weir B (2014) Letter to Editor in response to Editorial by Risinger et al. *Science and Justice* 54(6): 510.

Champod C (2014a) DNA transfer: Informed judgment or mere guesswork? In Biedermann A, Vuille J and Taroni F (eds) *DNA, Statistics and the Law: A Cross-Disciplinary Approach to Forensic Inference*. Lausanne: Frontiers Media S.A.

Champod C (2014b) Research focused mainly on bias will paralyse forensic science. *Science and Justice* 54(2): 107–109.

Cole S (2009) Forensics without uniqueness, conclusions without individualization: The new epistemology of forensic identification. *Law, Probability and Risk* 8(3): 233–255.

Cole S (2014) Individualization is dead, long live individualization! Reforms of reporting practices for fingerprint analysis in the United States. *Law, Probability and Risk* 13: 117–150.

Cook R, Evett I, Jackson G, Jones P and Lambert J (1998a) A hierarchy of propositions: Deciding which level to address in casework. *Science and Justice* 38(3): 231–240.

Cook R, Evett I, Jackson G, Jones P and Lambert J (1998b) A model for case assessment and interpretation. *Science & Justice* 38(3): 151–156.

Curran J, Hicks T and Buckleton J (2000) *Forensic Interpretation of Glass Evidence*. Boca Raton, FL: CRC Press.

European Network of Forensic Science Institutes (ENFSI) (2010) *Statement Regarding Interpretation*. BRD-GEN-004.

Evett I (1996) Expert evidence and forensic misconceptions of the nature of exact science. *Science and Justice* 36(2): 118–122.

Evett I and Weir B (1998) *Interpreting DNA Evidence*. Sunderland, MA: Sinauer Associates Inc.

Evett I, Jackson G and Lambert J (2000) More on the hierarchy of propositions: Exploring the distinction between explanations and propositions. *Science and Justice* 40(1): 3–10.

Finkelstein M and Fairley W (1970) A Bayesian approach to identification evidence. *Harvard Law Review* 83(3): 489–517.

Gallidabino M, Biedermann A and Taroni F (2015) Commentary on: Gauriot R, Gunaratnam L, Morini R, Reinikainen T, Corander R: Statistical Challenges in the Quantification of Gunshot Residue. *Evidence. J Forensic Sci* 2013;58 (5);1149–55. *Journal of Forensic Sciences* 60: 539–541.

Jackson G (2014) The impact of commercialization on the evaluation of DNA evidence. In Biedermann A, Vuille J and Taroni F (eds) *DNA, Statistics and the Law: A Cross-Disciplinary Approach to Forensic Inference*. Lausanne: Frontiers Media.

Jackson G and Jones P (2009) Case assessment and interpretation. In Jamieson A and Moenssens A (eds) *Wiley Encyclopedia of Forensic Science*. Chichester: John Wiley & Sons.

Jackson G, Jones S, Booth G, Champod C and Evett I (2006) The nature of forensic science opinion: A possible framework to guide thinking and practice in investigations and in court proceedings. *Science and Justice* 46(1): 33–44.

Jackson G, Aitken C and Roberts P (2013) *Case Assessment and Interpretation of Expert Evidence (Practitioner Guide No. 4), Guidance for Judges, Lawyers, Forensic Scientists and Expert Witnesses*. Royal Statistical Society's Working Group on Statistics and the Law. Available at: www.rss.org.uk/Images/PDF/influencing-change/rss-case-assessment-interpretation-expert-evidence.pdf (accessed 24 October 2016).

Lad F (1996) *Operational Subjective Statistical Methods: A Mathematical, Philosophical, and Historical Introduction*. Chichester: John Wiley & Sons.

Lindley D (1985) *Making Decisions*. Chichester: John Wiley & Sons.

Lindley D (1991) Probability. In Aitken C and Stoney D (eds) *The Use of Statistics in Forensic Science*. New York: Ellis Horwood.

Lindley D (2000) The philosophy of statistics. *The Statistician* 49(3): 293–337.

Lindley D (2014) *Understanding Uncertainty*. Chichester: John Wiley & Sons.

Neumann C, Evett I and Skerrett J (2012) Quantifying the weight of evidence from a fingerprint comparison: A new paradigm. *Journal of the Royal Statistical Society, Series A* 175(2): 371–415.

Quatrehomme G, Biglia E, Padovani B, du Jardin P and Alunni V (2014) Positive identification by x-rays bone trabeculae comparison. *Forensic Science International* 245: e11–e14.

Risinger M, Thompson W, Jamieson A, Koppl R, Kornfield I, Krane D, Mnooking J, Rosenthal R, Saks M and Zabell S (2014) Letter to the Editor: Regarding Champod, editorial: 'Research focused mainly on bias will paralyse forensic science'. *Science and Justice* 54(6): 508–9.

Robertson B and Vignaux A (1995) *Interpreting Evidence: Evaluating Forensic Science in the Courtroom*. Chichester: John Wiley & Sons.

Robertson B and Vignaux A (1998) Explaining evidence logically. *New Law Journal, Expert Witness Supplement* 148: 159.

Saks MJ and Koehler JJ (2005) The Coming Paradigm Shift in Forensic Identification Science. *Science* 309: 892–895.

Simmross U (2014) Appraisal of scientific evidence in criminal justice systems: On winds of change and coexisting formats. *Law, Probability and Risk* 13(2): 105–115.

Stoney D (1991) What made us ever think we could individualize using statistics? *Journal of the Forensic Science Society* 31(2): 197–199.

Taroni F, Champod C and Margot P (1998) Forerunners of Bayesianism in early forensic science. *Jurimetrics Journal* 38: 183–200.

Taroni F, Biedermann A, Vuille J and Morling N (2013) Whose DNA is this? How relevant a question? (A note for forensic scientists). *Forensic Science International: Genetics* 7(4): 467–470.

Willis S (2014) Accreditation—Straight belt or life jacket? *Science & Justice* 54(6): 505–507.

Appendix

Project Core Group: Sheila Willis (Project Leader, Forensic Science Ireland (FSI), Ireland); Louise McKenna, Sean McDermott, Geraldine O'Donnell (Partner, FSI, Ireland); Aurélie Barrett (Partner, National Institute of Criminalistics and Criminology (INCC), Belgium); Birgitta Rasmusson, Tobias Höglund (replaced by Anders Nordgaard) (Partners, National Forensic Centre (NFC), Sweden); Charles Berger, Marjan Sierps (Partners, Netherlands Forensic Institute (NFI), Netherlands); José Juan Lucena Molina (Partner, Criminalistic Service of the Civil Guard, Spain); Grzegorz Zadora (Partner, Institute of Forensic Research (IFR), Poland); Colin Aitken (Partner, University of Edinburgh, Scotland); Tina Lovelock (replaced by Luan Lunt) (Partner, LGCForensics, UK); Christophe Champod, Alex Biedermann, Tacha Hicks, Franco Taroni (Consultants, University of Lausanne, Switzerland).