

Noname manuscript No.
(will be inserted by the editor)

Normative decision analysis in forensic science

Received: date / Accepted: date

Abstract This paper focuses on the normative analysis – in the sense of the classic decision-theoretic formulation – of decision problems that arise in connection with forensic expert reporting. We distinguish this analytical account from other common types of decision analyses, such as descriptive approaches. While decision theory is, since several decades, an extensively discussed topic in legal literature, its use in forensic science is more recent, and with an emphasis on goals such as the analysis of the logical structure of forensic expert conclusions regarding, for example, propositions of common source of evidential and known materials. Typical examples are so-called identification (or, individualization) decisions, especially categorical conclusions according to which fingermarks (or stains of biological nature, handwriting, etc.) come from a particular a person of interest. We will present and compare ways of stating forensic identification decisions in decision-theoretic terms and explain their underlying rationale. In particular, we will emphasize the importance of viewing this analysis as normative in the sense of providing a reflective rather than a prescriptive reference point against which people in charge of forensic identification decisions may compare their otherwise (possibly) intuitive and informal reasoning, before acting. Normative decision analysis in forensic science thus provides a vector through which current practice can be articulated, scrutinized and rethought.

Keywords Normative decision analysis · Decision theory · Value of consequences · Forensic decision · Forensic expert reporting

A. Biedermann, S. Bozza, F. Taroni
University of Lausanne, Faculty of Law, Criminal Justice and Public Administration, School of Criminal Justice, 1015 Lausanne-Dorigny, Switzerland
E-mail: {alex.biedermann, silvia.bozza, franco.taroni}@unil.ch

S. Bozza
Ca' Foscari University of Venice, Department of Economics, 30121 Venice, Italy
E-mail: silvia.bozza@unive.it

“Even if you don’t analyze your decision problem by the methodology described in these lectures, you still must act. What will you do?” (Raiffa, 1968, p. 272)

1 Introduction

Evidence and decision-making are prevalent topics in legal scholarship, with a longstanding history, diverse methodological perspectives and concurrent theoretical accounts.¹ Despite the multitude of approaches covering normative, descriptive, prescriptive and interpretative views, certain concepts centering around decision-making in the law and adjacent areas, in particular criteria or rules of decision, remain controversial (e.g., Allen and Pardo, 2018). This poses a challenge to interdisciplinary exchange. It is useful, thus, to remind that in the context of law (Anderson et al, 2005), as well as in fact analysis in general (Schum, 1994), it has been argued that a discourse on evidence should include a clear indication of one’s standpoint. We shall do so at this point and emphasise that our background is not that of lawyers, but that of researchers and practitioners in forensic science (AB and FT) and statistics (SB). Our focus of attention is not legal decision (e.g., at trial, regarding ultimate issues), but questions of decisions as they may be encountered by forensic scientists. One of the viewpoints which we intend to substantiate is that many operational questions faced by forensic scientists, questions such as ‘How to conclude?’ or ‘What to report?’, can be understood as decisions, and hence be looked at from and analysed using formal theories of decision. Forensic decision questions, however, arise only at selected instances of the legal process, with no claim of generality for other aspects or parts of the trial process.

The structure of the paper is as follows. Section 2 provides a general statement of the formal theory of decision that we will mainly consider throughout this paper. We will also delimit the scope of our enquiry to forensic science. In Section 3 we will explain the sense in which we understand this formal theory of decision as normative. Further, we will distinguish normative decision analysis from other types of decision analyses (e.g., descriptive and prescriptive approaches). We will continue, in Section 4, with a critical analysis and comparison of two examples of standard rules of decision. The first one is

¹ An illustrative example for this is the recent ICAIL 2017 workshop *Evidence & Decision Making in the Law – Theoretical, Computational and Empirical Approaches* (The 16th International Conference on Artificial Intelligence and Law, June 16 2017, King’s College, London) which covered contributions to help make progress in our understanding of the following topics:

1. Modeling evidential reasoning and decision making at trial (e.g. evidence weighing; conflict resolution; standards of proof and rules of decision);
2. Evidence-based decision making and the architecture of the trial system (e.g. rules of admissibility; discovery procedures; adversary v. inquisitorial models; rules of weight v. free proof); and
3. The role and limitations of expected utility theory, and more generally cost/benefit analysis, for evidence-based decision making.”*

*<https://icail2017evidencedecision.wordpress.com/cfp/> (last accessed Aug 31 2018)

the result of a non-probabilistic approach. The second one involves probability and is based on expected value theory.² It will be seen that the latter is capable of providing a more comprehensive account than the former of the various ingredients that practical decision problems entail. The last section of the paper, Section 5, will discuss the value and practical impact of probabilistic decision theory (i.e., expected value theory) in forensic science and draw parallels to legal principles (e.g., the notion of ‘balance of probabilities’). The paper is didactic in Sections 2 to 4, providing an overview of perspectives and more detailed developments previously presented in Biedermann et al (2008, 2014, 2016a) and Taroni et al (2010). Discussions presented in Sections 5 and 6 provide a position statement on the extent to which decision-theoretic insight might meaningfully contribute to ongoing reforms of forensic reporting practice and the interface between forensic science and legal practitioners.

2 Decision theory and delimitation of the scope of enquiry to forensic science

The formal theory of decision that we consider for the most part in this paper emerged around the early 1950s. At that time, discussions arose about how decisions involving monetary consequences ought to be made in a sensible way, but also about behavioural aspects of how people actually make decisions (Lindley, 1985; Luce and Raiffa, 1958; Raiffa, 1968; Raiffa and Schlaifer, 1961; Savage, 1951, 1972; Smith, 1988). Central at the time was the theory of utility constructed by von Neumann and Morgenstern (1944). They provided conditions for preferences over decisions, thought of as rationality requirements, to be representable by a utility function.

In the same period, decision theory also stimulated interest among legal scholars, recognising it as a framework to formalise the question of how to qualify and compare decisions whose consequences are not completely known. Kaplan (1968), widely considered as being the first writer on the topic from a legal perspective, provided the following concise statement:

“Put in its most abstract form, the typical decision-theory problem involves the proper course of action to be taken by a decisionmaker who may gain or lose by taking action upon certain data that inconclusively support or discredit differing hypotheses about the state of the real but nonetheless unknowable world.” [at p. 1065]

Kaplan (1968) clearly recognised that formalisation is challenging and that it is important to distinguish the value of decision theory as an analytical framework from the theory’s immediate applicability at the time. He remained careful in his approach by emphasising that

² We use here the term expected value theory in a generic sense, as a synonym for classic decision theory, introduced in a non-technical way in Section 2. It is understood that the term ‘value’ may be replaced, for example, by the notions of utility or loss, depending on the area of application.

“(...) although we are in most legal areas far away from a useable quantification of our problems, the effort of thinking through the abstract quantitative relations among different variables turns out to provide a host of insights into the workings of our legal system.” [at p. 1066]

A defining feature of decision theory is that it combines probabilities for uncertain states of affairs or, more generally speaking, competing versions of an event of interest, with a quantification of desirability and undesirability of decision consequences in terms of utilities or losses.

In the decades after Kaplan’s (1968) formative paper, interests extended to include empirical perspectives. Studies have been conducted to see how well the coherent combinations of probabilities and utilities/losses according to decision theory align with people’s actual, but unaided judgment and decision behaviour. One of the findings was that people’s declared viewpoints (e.g., regarding probability thresholds) vary broadly (see, e.g., Hastie (1993) for a review), and may mismatch the formal decision-theoretic results. This is a descriptive result, however, commented on further in Section 3, but does not invalidate the formal character of decision theory.

In-depth discussions have been led, and divergent opinions exchanged, on the adequacy and appropriateness of decision theory – especially expected utility/loss criteria – for modelling decision making at trial. For example, professors Allen (2003) and Kaye (1999) placed different emphasis on the analytical perspective (Kaye) as compared to the empirical nature of the problem in the first place, especially the nature of the legal process (Allen). The decision points they consider regard, as in Kaplan’s (1968) paper, an advanced stage of the process, namely verdicts in court, which is different from the more restricted scope of forensic science questions considered in this paper. Our focus here is on decisions that may follow results of forensic examinations, such as ‘whose handwriting is this?’ or ‘who left this fingermark?’ Such questions, also known as identification or individualization³ questions (Champod, 2000), are not only of interest at selected instances in criminal cases, but also in civil litigation cases (e.g., involving signed contracts). In essence, these questions amount to asking ‘what to conclude?’ or ‘what to report?’ One of the viewpoints that we reiterate in this discussion paper is that such questions can be understood as decisions, and hence looked at through the formal theory of decision (Biedermann et al, 2008, 2016a; Taroni et al, 2010).

3 The notion of normative decision analysis : distinction with respect to other types of decision analysis

In discourses about decision analysis it is helpful and clarifying to distinguish between three main perspectives. They are referred to – especially in the field

³ In forensic science, the term individualization is commonly understood as the reduction of a pool of potential donors of a forensic trace (e.g., an impression, a mark or physical matter) to a single source, which may be an object or a person (Champod, 2000).

'judgment and decision making' of applied psychology (e.g., Baron, 2008) – as normative, descriptive and prescriptive. Descriptive decision analysis encompasses research on people's observable judgment and decision behaviour, in particular the development of psychological theories to explain behavioural patterns. This perspective also includes research on departures from reference or standard ways of behaviour (e.g., bias). The prescriptive approach, in turn, concentrates on ways to avoid such departures. It focuses on interventional measures to help improve decision behaviour, that is empowering people so that they become more resistant against undesirable ways of judgment and decision making. Besides psychology, the prescriptive approach thus also extends to the domains of engineering and design (Baron, 2012).

However, descriptive and prescriptive perspectives leave an essential type of question unanswered. It is the question of what is, in a particular situation, a desirable decision behaviour (e.g., rationality). Answers to this type of question are crucial for both the descriptive and the prescriptive approach. For example, to assess the 'goodness' of observed decision behaviour in the descriptive approach, a point of comparison is needed. Similarly, for the prescriptive approach, before thinking about how to improve people's decision behaviour (and associated decision strategies), it may be necessary to define which decision behaviour to consider desirable. These questions fall into the domain of normative decision analysis. Here, the term 'normative' is not understood as a legal norm (e.g., a procedural principle), but as branch of enquiry that focuses on philosophical and mathematical argument for (logical) standards of evaluation. An important feature of these standards is that they are justified independently of observed judgment and decision behaviour.

While the three approaches to decision analysis described above pursue different goals, their perspectives are not in competition, but complete each other. For example, the descriptive approach can only reveal recurrent errors of reasoning if a critical standard of comparison – as given by the normative perspective – is available in the first place. Stated otherwise, in order to expose a shortcoming in observable behaviour, it is necessary to define a point of comparison. While this reveals the crucial role that normative decision analysis can play, a recurrent and skeptical question is why a *theory* of decision should be looked at from a normative point of view. This question has a longstanding history. In the early days of decision theory, de Finetti – a sophisticated supporter of the normative perspective – addressed it as follows:

"A theory describing the behaviour of a decision-making agent (...) does not affect the way of deciding. The theory must be normative, (...) a requirement for consistency between decisions taken by an individual and the criteria he accepts (...)" (de Finetti, 1961, at p. 159)

De Finetti's argument here is in favour of coherence relationships between the fundamental ingredients of decision problems about which people's decision behaviour has little to say. Lindley, in his foreword to de Finetti's two-volume work *Theory of Probability* (1974), recently reprinted as de Finetti (2017), restated this position as follows:

“(...) experiments have been performed which show that individuals do not reason about uncertainty in the way described in these volumes. The experiments provide a descriptive view of man’s attitudes : de Finetti’s approach is normative. To spend too much time on description is unwise when a normative approach exists, for it is like asking people’s opinion of $2 + 2$, obtaining an average of 4.31 and announcing this to be the sum. It would be better to teach them arithmetic.” (de Finetti, 1974, at p. vii)

The normative approach, according to this account, emphasises that the formalised elements of a decision problem do not combine in arbitrary ways, but by particular rules that are not found by mere introspection or observation. Instead, they are derived through an analytical approach.

While the descriptive perspective is characterised by inherent limitations, normative accounts are challenged on their interpretation with respect to practical applications. In particular, in forensic science, it is relevant to enquire about and demonstrate how normative decision analysis can help practitioners approach decisional issues in their respective area of expertise (Taroni et al, 2005). Further, it is relevant to clarify the relative merit of different rules of decision following from different decision criteria. These topics are addressed in Section 4.

4 Critical analysis of two standard decision rules

4.1 A non-probabilistic decision rule

Instead of presenting decision theory as a mere definition, and in a programmatic way, it is helpful to lay out the formalisation of the forensic decision problem (or question), given at the end of Section 2, in a stepwise manner. We will proceed in two steps. First, we will provide the rudiments of a description of the decision problem and point out in what sense a given standard (non-probabilistic) decision rule (i.e., the minimax decision rule), in spite of its large appeal due to its ease of implementation, shows critical limitations. Second, in order to overcome the highlighted limitations, we will extend the basic description of the decision problem. This extension will lead us to formulate a probability-based decision criterion. Proceeding according to this staged approach will allow us to explain, at each step, the role of and need for each component of the decision-theoretic approach.

Start by considering the non-probabilistic decision rule through the following three elements: (i) actions (or, decisions) d_i , (ii) possible states of the world H_j , and (iii) decision consequences C_{ij} . Decision consequences define the outcome of choosing a particular action under a given state of the world. In a forensic decision problem, decision consequences may consist of the correct or incorrect acceptance of a hypothesis regarding, for example, the source of a handwritten document (e.g., whether a document was written by a given person of interest (POI) or by an unknown person). An example is shown in

		States of nature: The person of interest is ... the source (H_1) ... not the source (H_2)	
Decisions:	individualize (d_1)	C_{11}	C_{12}
	do not individualize (d_2)	C_{21}	C_{22}

Table 1 A simple forensic decision problem with two possible actions d_1 and d_2 , identifying or not a person of interest (POI) as the writer of a handwritten document, and two possible states of the world H_1 and H_2 , defined as ‘The POI is the author of the handwritten document’ and ‘An unknown person is the author’. In a general individualization problem ‘author’ may be replaced by ‘source’ and ‘handwritten document’ by ‘trace’ (mark, or impression). For each decision d_i , and each possible state of the world H_j , a consequence C_{ij} is obtained.

Table 1. The possible states of the world are whether or not the POI is the source of a given handwritten document; these possible states of nature represent competing versions of the events pertaining to the case at hand. The decisions are the acceptance or rejection of a particular version of the events.

In addition to the above three basic elements of a decision problem, a fourth element is introduced. It deals with the quantification of desirability or undesirability of decision consequences in terms of utilities or losses, expressing the relative worth, value, merit, loss or inconvenience as perceived or judged by the individual decision-maker. The quantification of utilities and losses is a broad topic in its own right; in some contexts, especially business applications, purely monetary values are used. For the purpose of the current discussion, we follow a perspective that avoids particular numerical assignments so as to favour broad acceptance. Specifically, in a first step, we will consider the following qualitative properties of preferences among decision consequences (see Table 1 for the definition of the consequences C_{ij}):

$$\underbrace{C_{12}}_{\text{‘worst’ consequence}} \prec C_{21} \prec \underbrace{C_{11} \sim C_{22}}_{\text{‘best’ consequences}}, \quad (1)$$

where \prec means ‘not preferred’. According to this ranking, we prefer correct identification (and non-identification⁴), C_{11} and C_{22} , to a false identification C_{12} . In a second step, it is possible to express this preference structure more formally, that is numerically, by characterising decision consequences C_{ij} in terms of, for example, their undesirability, using the concept of loss, written $L(C_{ij})$. There are different ways of choosing the origin of the numerical scale, as well as the range of possible loss values, though for ease of exposition it is useful to use the standard $0 - 1$ loss function. This loss function assigns the value 0 to the best (in the sense of less unfavourable or more favourable) consequence(s). Consequences that are judged adverse to some extent have a loss value greater than 0, with the value 1 assigned to the worst (less favourable or more unfavourable) consequence(s). The ranking (1) can thus be expressed

⁴ Non-identification is understood here not as ‘exclusion’, but as an unspecified alternative covering all situations *other* than ‘identification’.

with the following loss values:

$$L(C_{12}) = 1, \quad L(C_{21}) < 1, \quad L(C_{11}) = L(C_{22}) = 0. \quad (2)$$

Combining all the formal elements introduced so far, we can now state a standard (non-probabilistic) decision rule, referring to the *minimax decision principles* (e.g., Luce and Raiffa, 1958; Berger, 1985). It selects a decision on the sole basis of the specified (un-)desirability of decision consequences (here: losses). Specifically, an optimal decision can be found according to the following procedure: (a) highlight, for each decision the worst possible consequence, and (b) choose the decision that minimises the maximum loss over the different states of nature. Applied to our running example, step (a) highlights C_{12} (with loss 1) and C_{21} (with loss < 1) as the worst consequences of, respectively, decisions d_1 and d_2 . Step (b) then singles out d_2 as the optimal decision because the worst consequence that this decision may lead to, C_{21} , has a smaller loss than the worst consequence that may result from choosing d_1 (i.e., C_{12} with maximum loss 1, see also ranking (2)). Stated otherwise, the method described here thus says to look at the worst consequence(s) associated with each decision, and then choose the decision that implies the minimum loss if the worst case occurs (note that it is possible to have more than one decision that minimises the maximum loss).

In our application to forensic individualization, it quickly becomes apparent that such a decisional strategy has limitations, since it would mean to always decide d_2 , not individualising the POI, no matter how large the loss when an identification is missed, and no matter what information there may be in support of the state of nature θ_1 ('The POI is the source of the trace'). Proceeding in this way would ensure that, in case one's determination is erroneous, the loss incurred would be minimal (i.e., < 1). It is minimal with respect to the loss associated with the worst case (consequence C_{12}) of the alternative decision, d_1 , set to 1. Relying on such a criterion might seem attractive, because it strives to avoid excessive losses. But for forensic individualization, following this principle under the chosen loss function, would set an end to individualization, and hence paralyze current practice. Completely avoiding the possibility of incurring the worst consequence can only be achieved by never selecting the individualization decision d_1 . Overall, the minimax approach underlying this decision rule thus represents a pessimistic attitude, because of the systematic focus on the worst consequences.

Clearly, it is relevant and important for forensic science practice to be concerned with the worst consequence, i.e. a false identification. At the same time, forensic practitioners wish to maintain individualizations in the scope of their conclusions, and strive to pursue a framework in which individualization decisions can be articulated. There is merit, thus, in taking a closer look at the conditions under which an individualization decision d_1 might be reached, in order to see if these conditions can be captured in terms of a formal criterion. This directs our discussion to a standard probability-based decision rule.

4.2 Probabilistic decision rule (expected loss criterion)

Forensic scientists and other decision-makers in the legal process sometimes make the individualization decision (d_1), presumably when they hold a (sufficiently) high probability for the state of nature H_1 , the proposition according to which the POI is the source of the questioned handwriting, given all the relevant information I available. Denote this probability by $\Pr(H_1 | I)$.⁵ Stated otherwise, the forensic scientist who makes the individualization decision accepts the possibility of incurring the maximum loss $L(C_{12}) = 1$ when the probability $\Pr(H_2 | I)$ of the alternative state of nature H_2 , an unknown person being the source of the handwriting (i.e., on a document), given all the relevant information I , is (sufficiently) small. According to this view, making a decision is not exclusively based on only the consideration of the worst consequences that may incur under each decision, but by ‘keeping an eye’ on each decision consequence, as well on the probability of occurrence associated with each consequence. More formally, this can be understood as the weighting of the loss of each decision consequence by the probability of their occurrence. In mathematical terms, this amounts to computing the expectation (or, expected loss), defined as follows:

$$\text{EL}(d_i) = \sum_{j=\{1,2\}} L(C_{ij}) \Pr(H_j | I), \text{ for } i = \{1, 2\} \quad (3)$$

The expected loss characterises decisions d_i and provides a basis for their comparison, and hence the definition of a decision rule. The most well known expectation-based rule says to choose the decision that minimises expected loss (e.g., Lindley, 1985; Berger, 1985).

While the concept of expected value, and its computation, is standard material in decision science research, and essentially uncontroversial from a mathematical point of view, there is ongoing discussion – in various potential areas of application – about the feasibility and appropriateness of using expected value as a basis for choosing among decisions (e.g., Buchak, 2016). In forensic science, decision theory is not commonly used, currently. It has, however, the potential to provide original insights in some of the most fundamental topics of forensic science, such as individualisation, and is driving change in this area. The discussion in Section 5 will reflect on this, using examples from practice.

⁵ It is assumed here that this probability is also informed by any relevant evidence E bearing on H that may be available to the decision-maker at the time of the decision. Thus, one’s decision is preceded by inference based on available evidence. This inferential step, commonly dealt with probabilistically (Aitken and Taroni, 2004), is not covered here, and $\Pr(H_1 | E, I)$ is written more shortly as $\Pr(H_1 | I)$.

5 Value and practical impact of decision theory in forensic science

5.1 Descriptive capacity

It is important to emphasize again that decision theory, and the notion of expected value, is normative⁶ in the sense of providing a standard for reasoning and coherent decision. The concept, thus, does not claim to capture people's intuitive judgmental behaviour (Section 3). It follows from this that any empirical demonstration that people, *quantitatively*, do not reason so as to minimize/maximise expected value (e.g., maximising expected utility), or that they do not find the theory intuitive, is irrelevant to the tenets of the theory itself. Departures of unaided reasoning from normative results is not only little surprising, but actually foreseeable, because a computational step is involved, i.e. the calculation of a sum of products (i.e., Equation (3)).⁷ Given these observations, a legitimate question then is what value, if any, the normative account of decision might have. Throughout Sections 5.1 and 5.2, this question will be approached from different perspectives. Concluding remarks on counterarguments will be presented in Section 6.

Regarding the above-mentioned limitation with respect to description, it is worth noting that it concerns observable behaviour – not decision making aspirations. In fact, as we will elaborate further below, decision theory can be considered to have descriptive potential, at least in a *qualitative* sense, with respect to a decision-maker's goals. Stated otherwise, decision theory allows one to provide a formal criterion that reflects the idea that one does not only decide based on what one prefers (i.e., preferences among decision consequences), but also based on what one believes is actually the case (i.e., probabilities for the states of nature, informed by evidence). To illustrate this, consider the special case of a two-action decision problem with two possible states of nature, and the use of a $0 - k$ loss function, assigning a zero loss to the favorite or less undesirable consequence and a loss equal to $k > 0$ to the worst consequence (e.g. the loss function displayed in (2), where $k = 1$). For such a situation, the following decision criterion holds (e.g., Bernardo and Smith, 2000): decide d_1 if and only if:

$$\frac{\Pr(H_1 | I)}{\Pr(H_2 | I)} > \frac{L(C_{12})}{L(C_{21})}. \quad (4)$$

The ratio on the left in (4) expresses the decision-maker's beliefs about the possible states of nature at the time when a decision needs to be made.⁸ The term on the right expresses the loss ratio for adverse decision consequences. It

⁶ Note that the provision of a norm does not imply prescription. See also Section 5.2 on the notion of conditional advice.

⁷ Note that this is different for the minimax criterion mentioned in Section 4.1 which involves no computation, only a direct comparison between losses of decision consequences.

⁸ It is possible to understand this term as the prior odds or the posterior odds for the propositions of interest. The latter is the product of the prior odds and the likelihood ratio for any data considered relevant (e.g., Parmigiani, 2001).

is readily seen that this way of stating the conditions under which the inequality $EL(d_1) < EL(d_2)$ is satisfied amounts to a comparison between beliefs, on the one hand, and losses on the other hand. Specifically, the criterion (4) says that the more one finds a false identification (C_{12}) undesirable, compared to an erroneous non-identification (C_{21}), the higher should be one's belief in proposition H_1 , relative to proposition H_2 , in order for d_1 to be the optimal decision. It is interesting to note that this concisely expresses the common saying 'The more is at stake, the more you shall be sure (before you decide)', assuming that stakes refer to losses and 'being sure' refers to probability (i.e., beliefs about real world states). As such, the decision-theoretic criterion, exemplified in (4), thus captures, *descriptively*, some fundamental intuitive and generic understandings about the problem of how to decide under uncertainty: the idea of 'weighting' beliefs against relative losses. Equation (4) makes these intuitions formally precise, and provides a justification for them.

In a broader perspective, the formal theory considered here is also amenable to drawing parallels to legal principles, such as the precept 'deciding based on the balance of probabilities'. Interestingly, in view of the formal theory exposed here, this principle is warranted only under certain assumptions: i.e., when the ratio on the right in Equation (4) is one. In a generalized decision-theoretic reformulation, the 'balance of probability' standard would extend to, for example, 'the balance between probabilities and losses for adverse decision consequences' or 'the balance of probability moderated by the losses for adverse decision consequences'. The inherent mechanical perspective of the latter formulation can further be illustrated by translating the formulaic result (4) in graphical terms of a physical law, such as Archimedes' Law of Lever (Biedermann et al, 2016a). This law states that an equilibrium is established between two magnitudes if their distances are, with respect to a point of suspension, reciprocally proportional to their weights (e.g., Dijksterhuis et al, 1987). As shown in Figure 1(i), for example, two equal magnitudes A and B are at equilibrium if the lengths M and N are the same. Figure 1(ii) shows that if the magnitude B is greater than A, it is necessary to increase the length M, with respect to N, in order to establish an equilibrium: specifically, if B is twice as 'heavy' as A, the length M needs to be the double of N. More generally, the law says $A \times M = B \times N$, which can be rewritten as

$$\frac{M}{N} = \frac{B}{A}. \quad (5)$$

This equation states that the reciprocal of the ratio of the distances M and N is equal to the ratio of the two magnitudes A and B. It is readily seen that (5) has the same structure as (4), so that the length of the levers can be interpreted as the decision-maker's probabilities for the states of nature H_j , and the magnitudes as the losses associated with adverse decision consequences. This is shown in Figure 1(iii),⁹ recalling also that the product of losses (mag-

⁹ It is also possible to interpret the magnitudes as the probabilities and the length of the levers as the losses of adverse decision consequences. See Biedermann et al (2016a) for further discussion.

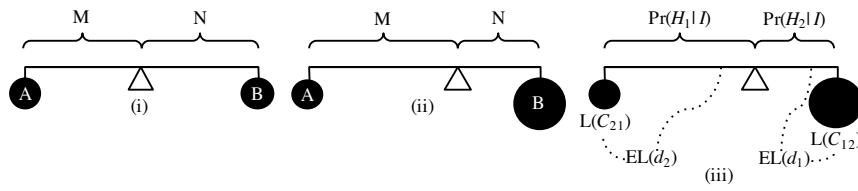


Fig. 1 (i) and (ii): Archimedes' Law of Lever illustrated with two magnitudes A and B at distances M and N from a pivot. (iii): Interpretation of Archimedes' Law of Lever in terms of the probabilities for propositions H_1 and H_2 (for the lengths of the distances) and the losses L (for the magnitudes) of wrong decisions. See Table 1 for definitions with respect to forensic individualization.

nitudes) and probabilities (length of levers) gives the expected losses EL of decisions d_i . Figure 1(iii) also illustrates that, for example, the more an erroneous identification C_{12} is considered worse – in terms of losses $L(\cdot)$ – than an erroneous non-identification, C_{21} , the greater the probability $\Pr(H_1 | I)$ must be compared to $\Pr(H_2 | I)$ (see Table 1 for a definition of the decision consequences C), in order for d_1 to be the decision with minimum expected loss. A sentence attributed to Archimedes is ‘Give me a place to stand, and I shall move the earth’ (e.g., Chondros, 2010; Dijksterhuis et al, 1987). In the context of the discussion of this paper, a possible analogy could be: ‘Whatever the weight (i.e., loss associated with) an erroneous identification, it can be levered (or, ‘outweighed’) if the probability for the hypothesis H_1 is sufficiently high (i.e., the lever is sufficiently long).’

Conversely, the Law of Lever also illustrates that even though the probability for the first hypothesis, H_1 , may be preponderant (i.e., much greater than the probability of H_2), it may be counterbalanced by a sufficiently large loss associated with C_{12} (a wrong individualization). That is, even though the probability $\Pr(H_1 | I)$ may be high, if the loss $L(C_{12})$ of a false identification is large enough, its multiplication with the probability of the alternative hypothesis $\Pr(H_2 | I)$ may lead to an expected loss $EL(d_1)$ that is *larger* than the expected loss of d_2 , hence making d_1 *not* the optimal decision.

The use of scales is common in legal discussions to illustrate, using a coarse language, the idea of ‘weighing’ evidence or probabilities – though such discourses assume that the scales typically have arms of equal and fixed length. The above analysis using Archimedes’ Law of Lever refines this intuition and makes it formally more precise, by highlighting a connection between a physical law and a conceptual decision criterion (i.e., the minimisation of expected losses). The analysis shows that decision is not only based on a comparison of probabilities, but on their weighing against the relative losses associated with the two ways in which one may decide erroneously. As a last remark, it is crucially important to insist on the point that the analysis focuses on decision making in the individual case. This means, in particular, that the relative adversity of decision consequences is assessed for the case at hand. There is no

reference to relative frequencies of errors of different kind assessed across *multiple* cases of individual decision making, a topic that relates to social policy (Kotsoglu, 2013).

5.2 The decisionalization of identification: decision-theoretic insight as a conditional advice

The viewpoints discussed in Section 5.1 may be considered controversial if they are understood as a suggestion that forensic identification, that is the decision to identify, a perspective also known as the decisionalization of identification (Biedermann et al, 2016b), reduces (or ought to reduce) to a purely mathematical question. This is not, however, the intended meaning and use of the analysis. The intention of the formal analysis is not to *delegate* (forensic) decision to an abstract mathematical theory. The theory merely provides every decision-maker with the analytical tools to help them clarify their thinking about relevant aspects of the decision problem, *before acting*. The latter two words, ‘before acting’, are crucial in this context, as further discussed below. Interestingly, this critical understanding of the theory as an analytical device has been clarified already by the founding writers on decision theory (e.g., de Finetti) – though this understanding is not always properly acknowledged in current treatments of this topic. De Finetti insisted on the understanding of decision-theoretic insight as a conditional advice, and clearly guarded against prescription:

“(...) it is not a question of prescribing anything, but only of giving each person a tool to enable them to better analyze and compare the pros and cons of each decision according to their own point of view, before choosing.” (de Finetti, 1961, p. 79)

The expression ‘before choosing’ should be emphasized because it conveys the idea that the role of decision theory is to provide an analytical framework for decision-makers to think carefully about a given decision problem, according to their personal point of view, and compare the various available courses of action based on a formal analysis. The idea is to empower individual decision-makers, by contributing to their broad scope of conceptual instruments.¹⁰

A parallel to this perspective can also be found in the positions of legal scholars and their understanding of the operational situation in which decision-makers in the legal process find themselves. Professor Allen, for example, has noted:

“Fact finders come to trial with a vast storehouse of knowledge, beliefs, and modes of reasoning that are necessary to permit communication to

¹⁰ The value of empowering individual decision-makers in various societal positions and functions is also recognised in areas beyond the law. An example for this in philosophy of politics is Mondadori who has been quoted as saying “[w]e can say that – given the importance of decisions in one’s every day life – it is a matter of democracy to have the opportunity to be supported by logical mechanisms for formal analysis of information.” (D’Agostino et al, 2001, p. 11)

occur simply and efficiently. (...) Everyone is just assumed to engage in orderly reasoning, employing all the necessary forms – deductive, inductive, abductive, statistical – as necessary or appropriate.” (Allen, 2015, p. 79)

The above reflections exemplify what is now recognised as a distinction between structural and interpretative elements (Buchak, 2016). In the formal theory of decision considered in this discussion (Section 4), the structural elements are the chosen function (here: a loss function) for quantifying the inconvenience associated with decision consequences, and the probability function for specifying beliefs about the possible states of nature. The structural element also covers a specified relationship between these two ingredients, here as defined by Equation (3), and its minimization. The interpretative element, in turn, pertains to the question of how the structural elements ought to be understood with respect to a particular area of application: there is a subtle difference between, on the one hand, saying that the theory can capture essential conceptual features of a decision-maker’s problem of making a well reflected choice and, on the other hand, actually arguing (i) that a decision – in any particular instance – ought to be made according to a criterion such as expected loss minimization, or (ii) that such a criterion defines rationality. Many debates on the feasibility and the merit of formalization in the legal process, such as the use of probability theory for measuring uncertainty and revising beliefs, gravitate around these interpretative questions. De Finetti’s defence of the theory as a conditional advice advocates a prudent position. He emphasises the theory’s capacity of unravelling the logical structure of decision problems, but carefully acknowledges that it does not actually prescribe particular choices. In Section 5.3 we will argue that this perspective also emerges from recent developments in forensic reporting practice.

5.3 The practical impact of decision-theoretic insight on current forensic reporting practice

Forensic practitioners who work in areas where so-called identification (or, individualization) conclusions are commonly reached now refer, increasingly often, to such conclusions as decisions, an understanding also noted by legal scholars (e.g., Broeders, 2009). An example for this is the title of the document ‘Guideline for the Articulation of the Decision-Making Process for the Individualization in Friction Ridge Examination’ (SWGFAST, 2013), issued by the Scientific Working Group of Friction Ridge Analysis, Study and Technology.^{11,12} Besides the document’s title, its Section 10.3.2 mentions “Decision-

¹¹ Note that SWGFAST evolved to what currently is the Subcommittee on Friction Ridge, which is part of the Organization of Scientific Area Committees (OSAC), administered by National Institute of Standards and Technology (NIST).

¹² The term ‘decision’ is also systematically used in a recent report issued by the AAAS (Thompson et al, 2017).

making in forensic identification" (referencing literature on decision theory), and Section 3.1 stipulates (emphasis as in original):

"Traditionally (for over 100 years) conclusions of individualizations were expressed differently, as an absolute identification: that *this person* did, *in fact*, make this impression, to the exclusion of all others in the world. As the practices of forensic science and of friction ridge examination have evolved, it is now recognized that our conclusions are more appropriately expressed as a decision, rather than proof. This decision is based on the support our findings provide to alternative possibilities (competing hypotheses)."

This statement is a moderate version of what Stoney has previously expressed more forcefully as follows:

"For over 100 years the courts and the public have expected, and fingerprint examiners have provided, expert testimony that fuses these three elements : offering testimony not as evidence, but as proof, assuming priors and including decision making preferences. This created an overwhelming and unrealistic burden, asking fingerprint examiners, in the name of science, for something that science cannot provide. As a necessary consequence, fingerprint examiners became unscientific." (Stoney, 2012, p. 400)

Interestingly, what SWGFAST discretely calls an 'evolution of practices' is revealed by Stoney as decision-theoretic insight. According to Stoney, analysing traditional identification conclusions using decision theory reveals that, tacitly, whoever makes a decision *must* admit a series of assumptions, such as evaluations of preferences among decision consequences, and beliefs about competing propositions of interest (i.e., the propositions whether or not a POI or an unknown person is the source of a given stain or mark). However, since such assumptions go above and beyond a scientist's area of competence, this type of traditional forensic decision practice is considered as having no scientific basis.

This does not mean, however, that SWGFAST and the wider community of forensic practitioners endorse decision theory, or define identification decisions in decision-theoretic terms, as shown by opinions collected by Cole (2014). Notwithstanding, there is now a growing movement towards thinking more in terms of the value of findings,¹³ rather than expressing direct opinions about hypotheses (e.g., that a given POI is the source of a trace or mark). This is a readjustment of practice, away from tacit assumptions about crucial ingredients of decisions, as noted by Stoney in the quote above, towards more defensible assertions about probative value. The following examples testify to the reality of this development:

¹³ The 'value of findings' is also sometimes referred to as the 'weight of evidence', that is an expression of the way and the extent to which particular (scientific) results help to discriminate between competing propositions of interest (e.g., Willis et al, 2015).

- SWGFAST refers to the notion of ‘value of findings’ with the sentence “the support our findings provide to alternative possibilities (competing hypotheses)” (SWGFAST, 2013, Par. 3.1). The two terms ‘support’ and ‘findings’ are central here. They define the area of competence of the scientist, whereas the hypotheses are merely the conditioning, taken as a given, on which no opinion will be expressed. How to define the notion of ‘support’ is a question in its own right, which goes beyond the scope of this paper (see, e.g., footnote 13 and further references given below).
- In 2015, the Defense Forensic Science Center (DFSC) of the U.S. Army Criminal Investigation Laboratory (USACIL) announced that it moves away from traditional ‘identification’ language to the following formulation (emphasis as in original):

“The latent print on Exhibit ## and the record finger/palm prints bearing the name XXXX have corresponding ridge detail. The likelihood of observing this amount of correspondence when two impressions are made by different sources is considered extremely low.”

(Swofford, 2015, Par. 2)

Clearly, these sentences express a probability¹⁴ for the findings *given* a proposition (rather than the contrary), which bears similarity with the components of the likelihood ratio. This measure of probative value is advocated, for example, in the guidance documents issued by ENFSI (Willis et al, 2015) and the Royal Statistical Society’s Working Group on Statistics and the Law (Aitken et al, 2010). Recently, the DFSC has issued a further information paper, reaffirming its previous position, but adding that it is now able to provide – to some extent – quantitative conclusions (see also Swofford et al (2018)). For example (emphasis as in original):

“The probability of observing this amount of correspondence is approximately ## times greater when impressions are made by the same source rather than by different sources.” (Swofford, 2017, Par. 2)

The above elements show that the contribution of decision theory to forensic science lies in the theory’s capacity to provide practising and academic discussants, and legal analysts, with a rigorous framework for critically examining current reporting practices. The theory also allows them to articulate the logical structure of decision-problems and to expose and explain deficiencies of current reporting formats. More generally, decision theory helps us argue that a change in forensic reporting practice is well in order, though the theory is silent about, hence does not directly prescribe, what exactly ought to be changed (see also discussion in Section 5.2). Some academics argue that as a consequence of decision-theoretic insight, in particular the exposed digression of forensic scientists above their area of competence (regarding probabilities and value

¹⁴ Note that the quote from the DFSC announcement considers the terms likelihood and probability as synonyms which, strictly speaking, is not correct from a statistical point of view.

judgments), “(...) experts should abandon the identification/individualization conclusion altogether.” (Champod et al, 2016, p. 96)

The restriction of current forensic reporting practice to probabilistic expressions of probative value, rather than asserted identification decisions, does not represent a loss or a failure. Rather, it helps ensure that reporting remains within defensible boundaries. Insofar, the role of decision-theoretic analysis is advisory: it provides a justification as to why probabilistic expressions of probative value sensibly define the scope conclusions that forensic scientists may report. Further elaboration of this perspective, in particular the call of Champod et al (2016) to abandon traditional reporting formats, is beyond the scope of this paper, but it is clear such decision-theoretic reflections contribute substantially to debates in ongoing reforms of forensic reporting practice.

6 Discussion and conclusions

Our analyses in this paper have concentrated on questions of decision making encountered in connection with forensic science results. It is important to keep in mind that such questions only concern a limited scope of issues that modern litigation systems face – a restriction acknowledged at the outset of this paper (Section 1). Clearly, there is a difference between reaching a verdict at the end of the legal process, taking into account the entirety of evidence provided, and addressing a selected problem of inference of source for which the only immediate specialised evidence is the one offered by the forensic scientist. Often, the latter decision point (e.g., in the context of forensic inference of source) is only one instance among many others contributing, though being distant from, the former question of decision.

The local nature of forensic decision problems offers some room for accommodating formalized analyses that otherwise, on the level of the legal process as a whole, encounter increased applicability problems (e.g., regarding the number of propositions and their assumed dependency relationships). Notwithstanding, some of the assumptions of the analytical decision-theoretic account studied here, in particular the specification of probabilities for states of nature and assignment of utilities and losses, pose challenges even on the level of idealised forensic decision problems. For example, it may be argued that forensic experts in legal contexts are not entitled – or are even discouraged – to maintain probabilities, and even less so utilities or losses. While this is an issue on a descriptive account, i.e. whether the theoretic account is capable of capturing the observable functioning of selected aspects of the legal process (i.e., decision making related to forensic evidence), it is crucial to enquire about the operational consequences of this observed mismatch. From our investigation of this topic we consider the most beneficial perspective to be the one advanced in Section 5.2, according to which the decision theory represents fundamental principles and results about decision making under uncertainty that are valuable for an informed decision-maker to know about *before* going into particular areas of application – especially when the case is

one in which intuitive one-off decisions should be avoided because the problem is intricate, the stakes are high (i.e., possibly severe consequences are involved) and enough time and resources for careful reflection are available. Although we understand decision theory in this context as normative, in the sense of providing a standard for coherent reasoning (Section 3), we consider decision-theoretic reflections as a conditional advice (Section 5.2), rather than as a rigid prescription. This ensures that decision-theoretic thinking remains compatible with further and more informal considerations that legal analysts count among the essential elements of decision making. As an aside, this formal perspective based on decision theory also reveals the particularities of decision problems in forensic and legal contexts that do not conform to traditional understandings in the broader spectrum of other human decisional activities.

Acknowledgements Alex Biedermann gratefully acknowledges the support of the Swiss National Science Foundation through grant No. BSSGI0_155809 and the University of Lausanne. Some of the writing of this paper was carried out during a visiting research stay of Alex Biedermann at New York University School of Law.

References

Aitken CGG, Taroni F (2004) *Statistics and the Evaluation of Evidence for Forensic Scientists*, 2nd edn. John Wiley & Sons, Chichester

Aitken CGG, Roberts P, 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. URL www.rss.org.uk/Images/PDF/influencing-change/rss-fundamentals-probability-statistical-evidence.pdf

Allen RJ (2003) The error of expected loss minimization. *Law, Probability and Risk* 2:1–7

Allen RJ (2015) A note to my philosophical friends about expertise and legal systems. *HumanaMente – Journal of Philosophical Studies* 28:71–86

Allen RJ, Pardo MS (2018) Relative plausibility and its critics. *Tech. Rep. 3179601*, The University of Alabama School of Law, URL <https://ssrn.com/abstract=3179601>

Anderson T, Schum DA, Twining W (2005) *Analysis of Evidence*, 2nd edn. *Law in Context*, Cambridge University Press, Cambridge

Baron J (2008) *Thinking and Deciding*, 4th edn. Cambridge University Press, New York

Baron J (2012) The point of normative models in judgment and decision making. *Frontiers in Psychology* 3, Article 577:1–3

Berger JO (1985) *Statistical Decision Theory and Bayesian Analysis*, 2nd edn. Springer, New York

Bernardo JM, Smith AFM (2000) *Bayesian Theory*, 2nd edn. John Wiley & Sons, Chichester

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

Biedermann A, Taroni F, Aitken C (2014) Liberties and constraints of the normative approach to evaluation and decision in forensic science: a discussion towards overcoming some common misconceptions. *Law, Probability and Risk* 13:181–191

Biedermann A, Bozza S, Taroni F (2016a) The decisionalization of individualization. *Forensic Science International* 266:29–38

Biedermann A, Bozza S, Taroni F, Aitken C (2016b) Reframing the debate: a question of probability, not of likelihood ratio. *Science & Justice, Special Issue on Measuring and Reporting the Precision of Forensic Likelihood Ratios* 56:392–396

Broeders T (2009) Decision-making in the forensic arena. In: Kaptein H, Prakken H, Verheij B (eds) *Legal Evidence and Proof, Statistics, Stories, Logic*, Ashgate Publishing, Farnham, pp 71–92

Buchak L (2016) Decision theory. In: Hájek A, Hitchcock C (eds) *Oxford Handbook of Probability and Philosophy*, Oxford University Press, Oxford, pp 789–814

Champod C (2000) Identification/individualisation, overview and meaning of ID. In: Siegel JH, Saukko PJ, Knupfer GC (eds) *Encyclopedia of Forensic Science*, Academic Press, San Diego, pp 1077–1084

Champod C, Lennard C, Margot P, Stoilovic M (2016) *Fingerprints and Other Ridge Skin Impressions*, 2nd edn. CRC Press, Boca Raton

Chondros TG (2010) Archimedes' influence in science and engineering. In: Paipetis S, Ceccarelli M (eds) *The Genius of Archimedes – 23 Centuries of Influence on Mathematics, Science and Engineering, History of Mechanism and Machine Science* 11, Springer, Dordrecht, pp 411–425

Cole SA (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

D'Agostino M, Giorello G, Veca S (2001) "Ragionare dobbiamo, e spesso". In: D'Agostino M, Giorello G, Veca S (eds) *Logica e politica, per Marco Mondadori, il Saggiatore*, Milano, pp 11–14

Dijksterhuis EJ, Dikshoorn C, Knorr WR (1987) *Archimedes*. Princeton University Press, Princeton, New Jersey

de Finetti B (1961) Dans quel sens la théorie de la décision est-elle et doit-elle être "normative". In: *Colloques Internationaux du Centre National de la Recherche Scientifique, "La Decision"*, Éditions du Centre National de la Recherche Scientifique, Paris, pp 159–169

de Finetti B (1974) *Theory of Probability, A Critical Introductory Treatment*, Volume 1. John Wiley & Sons, London

de Finetti B (2017) *Theory of Probability, A Critical Introductory Treatment*, reprint edn. John Wiley & Sons, Chichester

Hastie R (1993) Algebraic models of juror decision making. In: Hastie R (ed) *Inside the Juror. The Psychology of Juror Decision Making*, Cambridge

University Press, New York, pp 84–115

Kaplan J (1968) Decision theory and the factfinding process. *Stanford Law Review* 20:1065–1092

Kaye DH (1999) Clarifying the burden of persuasion: what Bayesian decision rules do and do not do. *The International Journal of Evidence & Proof* 3:1–29

Kotsoglou KN (2013) How to become an epistemic engineer: what shifts when we change the standard of proof? *Law, Probability and Risk* 12:275–298

Lindley DV (1985) *Making Decisions*, 2nd edn. John Wiley & Sons, Chichester

Luce RD, Raiffa H (1958) *Games and decisions : introduction and critical survey*. Wiley, New York

von Neumann J, Morgenstern O (1944) *Theory of Games and Economic Behavior*. John Wiley & Sons, New York

Parmigiani G (2001) Decision theory: Bayesian. In: Smelser N, Baltes P (eds) *International Encyclopedia of Social and Behavioral Sciences*, Elsevier, Oxford, p 33273334

Raiffa H (1968) *Decision Analysis, Introductory Lectures on Choices under Uncertainty*. Addison-Wesley, Reading, Massachusetts

Raiffa H, Schlaifer R (1961) *Applied Statistical Decision Theory*. The M.I.T. Press, Cambridge, Massachusetts

Savage LJ (1951) Theory of statistical decision. *Journal of the American Statistical Association* pp 55–67

Savage LJ (1972) *The Foundations of Statistics*, 2nd edn. Dover, New York

Schum DA (1994) *Evidential Foundations of Probabilistic Reasoning*. John Wiley & Sons, Inc., New York

Smith JQ (1988) *Decision analysis: a Bayesian approach*. Chapman and Hall, London

Stoney DA (2012) Discussion on the paper by Neumann, Evett and Skerrett. *Journal of the Royal Statistical Society: Series A (Statistics in Society)* 175:399–400

SWGFAST (2013) Guideline for the articulation of the decision-making process for the individualization in friction ridge examination, ver. 1.0. https://www.nist.gov/sites/default/files/documents/2016/10/26/swgfast_articulation_1.0_130427_1.pdf, last accessed 6 August 2017

Swofford H (2015) Use of the term “identification” in latent print technical reports. Information Paper, U.S. Army Criminal Investigation Laboratory, Defense Forensic Science Center, 3 November 2015

Swofford H (2017) Modification of latent print technical reports to include statistical calculations. Information Paper, U.S. Army Criminal Investigation Laboratory, Defense Forensic Science Center, 9 March 2017

Swofford HJ, Koertner AJ, Zemp F, Ausdemore M, Liu A, Salyards MJ (2018) A method for the statistical interpretation of friction ridge skin impression evidence: Method development and validation. *Forensic Science International* 287:113–126

Taroni F, Bozza S, Aitken CGG (2005) Decision analysis in forensic science. *Journal of Forensic Sciences* 50:894–905

Taroni F, Bozza S, Biedermann A, Garbolino G, Aitken CGG (2010) Data Analysis in Forensic Science: a Bayesian Decision Perspective. *Statistics in Practice*, John Wiley & Sons, Chichester

Thompson WC, Black J, Jain A, Kadane J (2017) Latent Fingerprint Examination. *Forensic Science Assessments: A Quality and Gap Analysis*, American Association for the Advancement of Science, Washington, D.C.

Willis S, McKenna L, McDermott S, O'Donell G, Barrett A, Rasmusson B, Nordgaard A, Berger C, Sjerps M, Lucena-Molina J, Zadora G, Aitken C, Lovelock T, Lunt L, Champod C, Biedermann A, Hicks T, Taroni F (2015) ENFSI guideline for evaluative reporting in forensic science, Strengthening the evaluation of forensic results across Europe (STEEOFRAE). Dublin