THE DECISIONAL NATURE OF PROBABILITY AND PLAUSIBILITY ASSESSMENTS IN JURIDICAL EVIDENCE AND PROOF

Alex Biedermann*, Joëlle Vuille**

* Faculty of Law, Criminal Justice and Public Administration, School of Criminal Justice, University of Lausanne, 1015 Lausanne-Dorigny (Switzerland)
** Faculty of Law, University of Fribourg, 1700 Fribourg (Switzerland)

ABSTRACT

The quantification of uncertainty is a key topic in different theories and accounts of the legal process, ranging from probabilism to explanationism. These accounts invoke probability to various extents. For legal probabilism, probability is the single core concept, whereas other accounts, such as the relative plausibility theory, give it a more limited role, as one consideration among others. At the same time, controversies persist about the nature of probability and the value it may add to the understanding of the broad range of aspects that characterise the legal process. These controversies arise, in part, from the fact that probability itself is the object of confusing debates in many scientific disciplines. In view of these intricacies, this paper argues that the critical analysis and clarification of how to understand and use probability meaningfully remains a topic worthy of investigation across different theoretical perspectives. The first part of this paper critically examines a selection of persisting misconceptions about probability and objections against its use, based on discussions presented in recent evidence law literature. Part II of this paper will blend this discussion with a particular view of probability, interpreted as a personal decision, rarely acknowledged in legal literature. Using a multidisciplinary perspective and a thorough review of historical sources, we illustrate and discuss how the understanding of probability assertion as a decision promotes transparency, honesty, accountability and justifiability. This decisional perspective is further developed and discussed in Part III to show that its logical ingredients underpin key concepts of different theoretical accounts, in particular the assertion of degrees of belief, assessments of relative plausibility and verdicts about ultimate issues at trial. Overall, the paper makes the point that the isolated debates over probability, legal probabilism and reasoning under uncertainty misconceive the primary problem of the legal process, which is decision-making under uncertainty. The proposed decisional perspective clarifies these issues both analytically and descriptively, and resolves divergencies between ostensibly competing concepts such as probability and relative plausibility.

INTRODUCTION

Doctrine and policy on evidentiary proof processes in criminal and civil cases, across modern legal systems, commonly invoke the understanding that decision makers[1] ought to consider the degree to which they are convinced

---

[1] Throughout this paper we will use 'decision makers' as a generic term in place of judges and jurors.

that particular legal elements are established, before making a decision. In a given case, decision makers need to address the question of whether they are sufficiently persuaded to decide in favour of a given party. The exact level of persuasion required – the standard of proof – applicable in a given case is an intricate evidentiary topic. While legal systems appear to leave it largely to decision makers to determine when a given level of proof is achieved in the case at hand, or a quantum of evidence was provided,[2] there is a broad scope of legislation, jurisprudence and academic writing on how this level of satisfaction should be reached. On the one hand, notions such as "preponderance of the evidence" (or "balance of probabilities") in civil cases, and "beyond a reasonable doubt" in criminal cases, are commonly explained in overtly probabilistic terms.[3] On the other hand, attempts by evidence scholars to provide a rigorously probabilistic account of standards of proof face a number of challenges and are the subject of decades-long debates.[4]

In a recent paper, Professors Allen and Pardo[5] review several accounts of the nature of juridical proof and argue in favour of a relative plausibility account (also called 'explanatory account'). Unlike approaches that consider probability as a starting point to capture selected aspects of the legal proof process, Allen and Pardo start from the litigation process as a whole, and how it structures proof at trial. In a second step, they consider how relative plausibility, empirically and descriptively, fits with the legal system's goals. Adherents of relative plausibility are not dismissive of probability, however,

---

[2] In some particular areas of application, such as paternity cases, it is common to refer to specific numerical thresholds.

[3]E.g., Brown v. Bowen, 847 F.2d, 342 (7th Cir. 1988) ("All burdens of persuasion deal with probabilities."), Ronald J. Allen & Alex Stein, Evidence, Probability, and the Burden of Proof, 55 Arizona Law Review 557–602 (2013), at 565 ("Burdens of proof easily bear a probabilistic interpretation."), and for references to old cases see also James P. McBaine, Burden of Proof: Degrees of Belief, 32 California Law Review 242–268 (1944). A pending question is, however, whether the use of the term 'probability' reflects the willingness to apply the calculus of probability or whether, as suggested by Haak, probability is understood merely as 'degree of warrant' (Susan Haak, Evidence Matters, Science, Proof, and Truth in the Law, 47–77, 2014).

[4] See, e.g., Peter Tillers & Eric D. Green (eds.), Probability and Inference in the Law of Evidence: The Uses and Limits of Bayesianism (1988), and the article collection edited by Ronald J. Allen & Mike Redmayne, Bayesianism in Juridical Proof, 1 International Journal of Evidence & Proof 253–360 (1997). More recently, see also Kevin M. Clermont, Standards of Decision in Law: Psychological and Logical Bases for the Standard of Proof, Here and Abroad (2013) and Dale A. Nance, The Burdens of Proof: Discriminatory Power, Weight of Evidence and Tenacity of Belief (2016). Historically, a hallmark in the critical review of evidence and proof accounts, especially legal probabilism that still echoes today, dates back to the early 1970 (Laurence H. Tribe, Trial by Mathematics: Precision and Ritual in the Legal Process, 84 Harvard Law Review 1329–1393, 1971).

[5] Ronald J. Allen & Michael S. Pardo, Relative Plausibility and Its Critics, 23 International Journal of Evidence & Proof 5–59 (2019).

because they consider probability to be one of the variables that inform judgments, among others aspects such as consistency, completeness and economy.[6] But if probability has to play a role in evidential assessment and juridical proof, as this view suggests, then the proper understanding of probability is essential. This is an intricate topic mainly because probability itself is a controversial topic across scientific disciplines. Thus, the critical analysis, discussion and, if possible, resolution, of objections recurrently levelled against legal probabilism is a topic worthy of investigation for both adherents of the relative plausibility theory of trial and probabilistic-centered proponents of the trial rationale.

In the first part of this paper, we critically review and discuss a selection of misconceptions of probability and objections against its use in legal settings observed in recent evidence law literature. Although several of these critiques have been formulated and known for a long time, the replies offered hitherto by evidence scholars and discussants from other domains, such as philosophy of science, appear to remain unconsidered or are felt to be unconvincing. In Part two of this paper, we seek to improve on this situation by building on de Finetti's operational and decision-oriented perspective to probability.[7] This is an unconventional account of probability and sceptical readers might immediately object to this framework as being incompatible with the values of the legal process, essentially because of the allegedly unconstrained subjectivist probability assignments that this perspective allows. We will explain in detail why this is a misconception of the personalist account of probability. Specifically, we will defend the view that the proper understanding of the personalist account actually encourages highly desirable values such as transparency, honesty, accountability and justifiability. We will substantiate this view with the help of a conceptual device that is based on the understanding of probability assertion as a *personal decision*, requiring careful reflection and the taking of responsibility. Further, we will argue that the debates over the suitability and nature of probability (objectivist, subjectivist or otherwise) in the legal setting distract from the principal issue and duty with which the legal process is concerned, which is decision under uncertainty. In Part three of this paper, we show that – conceptually – the notion of decision underpins key concepts ranging from the expression of degrees of belief in terms of either probability or plausibility, to verdicts at the end of the process. In this last part, we will

---

[6] Ronald J. Allen, The Nature of Juridical Proof: Probability as a Tool in Plausible Reasoning, 21 International Journal of Evidence & Proof 133–142 (2017) at 140, Ronald J. Allen, Rationality, Algorithms and Juridical Proof: A Preliminary Inquiry, 1 International Journal of Evidence & Proof 254–275 (1997) at 274.

[7] Bruno de Finetti, Theory of Probability: A Critical Introductory Treatment (1974; Reprint 2017).

discuss the analytical character of the proposed decisional perspective and distinguish it from its descriptive capacity. That is, our focus is not primarily on whether the formal approach to decision developed here provides a good *description* of how participants in the legal process actually behave. Instead, as we will argue, the decisional account captures the logical structure and inevitable ingredients of the decisional questions that any person required to act under uncertainty faces, irrespective of whether those actions are based on probability or on other concepts, as comprised in relative plausibility. As an aside, our analyses will also make the point that the mainstream debates over probability thresholds as bases for decisions are conceptually short-sighted and, in this regard, rightly criticized. The decisional account that we invoke clarifies why both, analytically and descriptively, sensible decision-making is not based on probability alone.

Methodologically, our paper is multidisciplinary. It critically examines the use of formal methods of decision analysis in the law, based on a careful interpretation of their historical sources. This will lead us to refer frequently to the writings of Professors Allen and Pardo because they are a primary source of serious challenges on this topic.

We insist on delineating this paper as focusing on probability, not on Bayes' theorem, and even less on Bayesianism. Pure probabilists may view these distinctions as artificial because Bayes' theorem is nothing but an inevitable consequence of the rules of probability. There is a difference, however, on a practical level. The full application of Bayes' theorem in the legal context – also referred to as 'judicial Bayesianism'[8] – faces a host of challenges, ranging from the specification of relevance relationships (i.e., conditionalization), to the assessment of component probabilities and the limitations in computational capacities at the level of detail that the theory stipulates. A moderate probabilistic account,[9] such as the one envisioned here, is primarily concerned with the coherent expression of uncertainties about target events in terms of probabilities. It is more readily defensible and admitted also by critics. As such, it is an important preliminary of judicial Bayesianism but also other approaches, such as relative plausibility, that

---

[8] Alex Stein, Judicial Fact-finding and the Bayesian Method: the Case for Deeper Scepticism About Their Combination, 1 International Journal of Evidence & Proof 25–47 (1997), at 25. An example of proponents of judicial Bayesianism is Bernard Robertson & George A. Vignaux, Probability – The Logic of the Law, 13 Oxford Journal of Legal Studies 457–478 (1993). For a defense of Bayesianism as an epistemology prescribing how one should answer the question of what to think or believe regarding particular propositions of interest, see Elliott Sober, Evidence and Evolution, The Logic Behind the Science (2008).

[9] Constructive and reconciling accounts of probability in the law can be found in the many writings of Professor Richard D. Friedman (e.g., Richard D. Friedman, The Persistence of the Probabilistic Perspective, 48 Seton Hall Law Review 1589–1600, 2018).

reserves at least a limited role to probability, among other considerations.[10] Thus, our concern here is only the nature and feasibility of probability assessments, regardless of whether they are direct assessments at the end point of an analysis, or component probability assessments that are either processed further using Bayes' theorem (by those who wish to take this additional step),[11] or used as parts of other accounts, such as relative plausibility.[12]

PART 1: MISCONCEPTIONS REGARDING THE NATURE OF PROBABILITY AND RECURRENT OBJECTIONS AGAINST ITS USE IN THE LEGAL SETTING

A. Probability: the problem of where to start

In a recent paper, Dahlman[13] addresses the issue of how a coherent legal reasoner ought to set initial probabilities – commonly called 'prior probabilities' – about competing propositions of interest advanced by opposing parties at trial. The terms 'initial' and 'prior' are often used interchangeably to clarify that the probabilities under consideration refer to the view of the decision maker before acquiring new evidence. The feasibility of conceptually defining and defensibly assigning such probabilities is widely considered a key factor on which hinges the suitability of probability for dealing with uncertainty in the legal process.[14]

The opening statement of Dahlman's paper provides a good example of a series of convoluted topics that require careful discussion: "A Bayesian assessment of the probability that the defendant in a criminal trial is guilty depends on the presumed base rate for guilt and the estimated likelihoods of the evidence."[15] To scrutinise the foundations of this statement, it is first necessary to recall the formal relationship between a proposition (or, hypothesis) *H* and an item of evidence *E*. Generally, the observation of an item of evidence *E* may strengthen, weaken or leave one's belief in the truth

---

[10] Supra note 6.

[11] E.g., De Finetti, Probabilism: A Critical Essay on the Theory of Probability and on the Value of Science, 31 Erkenntnis 169–223 (1989), at 196.

[12] Allen (2017), supra note 6.

[13] Christian Dahlman, Determining the Base Rate of Guilt, 17 Law, Probability & Risk 15–28 (2018). We mention this paper merely as an example. The assessment of initial (or so-called, prior) probabilities is a topic that also deeply perplexes scholars in areas other than the law, such as statistics and philosophy of science.

[14] This calls for reiteration of the conflicts that arise between properties of formal models and particular aspects of the justice process. This is illustrated, for example, by the controversies over the relationship between the specification of prior probabilities and legal presumptions, such as the presumption of innocence (e.g., David H. Kaye, The Probability of an Ultimate Issue: the Strange Case of Paternity Testing, 75 Iowa Law Review 75–109 (1989); Richard D. Friedman, A Presumption of Innocence, Not of Even Odds, 52 Stanford Law Review 873–887 (2000)).

[15] Id. at 15.

of the proposition *H* unaffected, depending on the probative value of *E*.[16] However, the diagnosticity of the evidence *E*, by itself, does not entitle us to make an assertion about the probability of the proposition *H*. Coherently working out a probability for the proposition *H* given evidence *E*, using the rules of the probability calculus, requires a prior probability of *H*, i.e. a probability without the conditioning on *E*. How to specify this probability in legal applications is the question that Dahlman strives to answer, although there are several complications in his attempt.

To start with, he refers to the prior probability as the base rate. This is a particular, but ill-conceived, definitional choice. As we will point out, it is at the heart of a host difficulties commonly ascribed to the probabilistic account of judicial proof.

The notion of base rate relates to contexts in which the particular case to be assessed is thought of within the broader picture of background incidences of events of a comparable kind. Most typically, in medical diagnosis, the assessment of the probability that a particular person has a given disease, once the result of a diagnostic test has been obtained, requires one to take into account the prevalence of the disease in the population of interest. This, in turn, raises a plethora of interrogations about whether, and if so, how, relevant populations ought to be defined. This is generally known, legal literature included, as the 'reference class problem'.[17,18] Difficulties increase – to the point of becoming unmanageable – whenever one considers highly particular combinations of target features, which is all too common in real legal cases.[19] While these are valid challenges that demonstrate a "deep conceptual problem"[20], it is important to recall that they are the consequence of a particular way to interpret and assess probabilities, and not – as we will argue throughout this paper – a problem of probability as such.

---

[16] In areas of application other than the law, probative value or strength of evidence is sometimes referred to as diagnostic or discriminative capacity. Probabilistically, the strength of the evidence depends on the probability of the evidence *E* if proposition *H* is true, compared to the probability of the evidence *E* given that proposition *H* is false. See, e.g., Colin Howson & Peter Urbach, Scientific Reasoning, The Bayesian Approach, Third Ed. (2006), Thomas D. Lyon & Jonathan J. Koehler, Relevance Ratio: Evaluating the Probative Value of Expert Testimony in Child Sexual Abuse Cases, 82 Cornell Law Review 43–78 (1996), Richard Friedman, The Elements of Evidence, Fourth Ed. (2017).

[17] E.g., Ronald J. Allen & Michael S. Pardo, The Problematic Value of Mathematical Models of Evidence, 36 The Journal of Legal Studies 107–140 (2007).

[18] For example, if the target characteristic is 'face tattoo', the standard challenge posed by the reference class problem is that there is a variety of ways to define the reference population from which data is to be collected in order to quantify the background incidence of people with tattoos on their face.

[19] For examples showing that the reference class problem can be easily pushed ad absurdum see, e.g., Allen (2017), supra note 6, at 136.

[20] Supra note 17, at 109.

This is not to say that base rates are not a relevant concept in a variety of settings, such as clinical practice.[21] It is relevant to inquire, however, whether they have any bearing on the inferences and decision problems encountered in legal cases. In its most general form the notion of base rate relates to the proportion of members of a given population that have a particular trait. It involves two ideas: (i) relative frequency, i.e. the number of times a particular observation is made, divided by the total number of observations, and (ii) long-run repetitions of particular experiments under the idealised assumption of stable (i.e. identical) conditions.[22] In the last decades, these ideas have been used in attempts to define and apply probability in general, and in the law in particular. The persistence of these attempts is rather surprising, especially in more recent literature, because there are several well-grounded criticisms, known for a long time and to date unrefuted. In the legal literature, for example, Kaplan noted:

> "Given a typical contested trial, for instance, it is meaningless to speak of the probability of the defendant's guilt in terms of the number of times he would be guilty in an infinite number of exactly similar cases because, first, there are not even two exactly similar cases, and, second, even if there were many identical cases the court must reach a verdict, not a ratio, in the case at bar."[23]

Similarly, Friedman observed: "Matters at issue in litigation tend to be one-time events. If Victor was murdered, it makes no sense to say, 'The probability that in any given murder of Victor the murderer would be Dennis is 1 in 2'."[24] The mathematical and statistical literature is in line with these views.[25]

---

[21] For an early recognition of the importance of base rates see, e.g., Paul E. Meehl & Albert Rosen, Antecedent Probability and the Efficiency of Psychometric Signs, Patterns, or Cutting Scores, 52 Psychological Bulletin 194–216 (1955).

[22] Alternatively, base rates are also sometimes thought of in terms of the tendency (also referred to as propensity) of a procedure, applied across multiple instances, to lead to a particular outcome at any given instance of application (see, e.g., Howson & Urbach, supra note 16, at 24–25).

[23] John Kaplan, Decision Theory and the Factfinding Process, 20 Stanford Law Review 1066 (1968).

[24] Richard D. Friedman, Answering the Bayesioskeptical Challenge, 1 The International Journal of Evidence & Proof 276–291 (1997) at 276. Similarly, Allen & Stein, supra note 3, argue that courts "(...) do not base their decisions on the frequencies of events that resemble the event they are trying to reconstruct (even when those frequencies are available)." (at 567)

[25] See, e.g., Dennis V. Lindley, Probability, in: Colin G.G. Aitken & David A. Stoney (Eds.), The Use of Statistics in Forensic Science (1991): "What is the chance that the defendant is guilty? Are we to imagine a sequence of trials in which the judgements, 'guilty' or 'not guilty', are made and the frequency of the former found? It will not work (...). The whole idea of chance is preposterous in this context" (at 48) and Bruno de Finetti, Philosophical Lectures on Probability, Alberto Mura (Ed.) (2008): "The frequency with

The assertion that "the probability that the defendant in a criminal trial is guilty depends on the presumed base rate for guilt"[26] thus is conceptually and definitionally unsound, but there are even more fundamental problems than trying to understand probability as a question of framing the relevant class or population, and the frequency of occurrence of selected features. Specifically, and we note this as an aside, it is unsound to refer to a 'probability of guilt' (or a "guilt-hypothesis"[27]) because guilt is not a hypothesis, but a legal conclusion.[28] This renders the attempt to conceive of a way to "determining the base rate rate of guilt"[29] pointless.

Sure, probabilistic inference about any hypothesis of interest requires a probability to start from. The rules of probability say that one cannot make an assertion about the probability of a hypothesis, given particular evidence, without stating what one's probability was *before* considering the evidence. The rules of probability do *not* require, however, one to determine the initial probability "according to some uniform recipe"[30], especially not in some frequentist way. What is more, thinking that one could elaborate principles and abstract procedures that claim to produce probabilities from scratch, that is without reference to particular information, is epistemologically objectionable.[31]

It follows from the above considerations that it is not meaningful to conceptualise a prior probability in the legal context in terms of a base rate. It might be objected that there are situations where base rates are available, such as in the following example: "Suppose, e.g. that a defendant was driving on a road where 60% of the drivers were intoxicated at the time when the defendant was stopped by the police. This would set the base rate for guilt at 0.60 (...)".[32] The critical point here is the transition from the first to the second of the previous two sentences, which is not one of necessity. Strictly speaking, the 60% figure is *data* only. It is to be distinguished from the *probability* that a particular individual is indeed drunk, which is a distinct

---

which certain events obtained or will obtain cannot be identified with probability. Frequency is a mere fact, independent of both the meaning of probability and the probability values assigned to the events" (at 138).

[26] Supra note 13, at 15, and similarly in Norman Fenton, David Lagnado, Christian Dahlman & Martin Neil, The Opportunity Prior: a Proof-Based Prior For Criminal Cases, Law, Probability & Risk (forthcoming, https://doi.org/10.1093/lpr/mgz007), ("assessing the probability that a suspect is guilty", at 2).

[27] Supra note 13, at 16.

[28] E.g., Allen (1997), supra note 6, at 265, and Allen (2017), supra note 6, at 138.

[29] Supra note 13, at 15.

[30] Howson & Urbach, supra note 16, at 25.

[31] See, for example, Wesley C. Salmon, The Foundations of Scientific Inference (1967) for a critique of the indifference principle, considering it "epistemological magic" (at 66).

[32] Supra note 13, at 21.

issue.[33] So, the question we are really interested in is the condition of a given individual or, to be more precise, our belief that this individual is in a particular state while having data on *other* individuals, supposedly comparable, in some sense, to the individual at hand. The key point is how to pin down this probability, because it does not follow directly from the available data. Lindley noted:

> "(...) you pass from a frequency among the patients seen, to a probability or belief about a further patient. This passage is so common that there has grown up a confusion between frequency, which refers to data, and probability, which is belief, so that people speak of the frequency interpretation of probability. There is a connection between the two concepts, but it is wrong to identify them (...)".[34,35]

Underlying this distinction is an alternative view of probability that interprets probabilities as measures of degrees of a person's belief, referred to hereafter as personal or subjective probabilities. This type of probability refers to a person's entirety of knowledge, rather than to a process that supposedly produces certain outcomes at a particular rate. The knowledge which conditions personal probabilities may, but need not necessarily, cover frequency data. It is thus possible to avoid the perceived inadequacies that derive from attempts to define and assess probabilities in terms of frequency data. In later sections of this paper, we will further build on and explain the fundamental difference between the definition and the assignment of probabilities. Although we recognize that personal probabilities face their own critiques, we will illustrate how to assess them in light of a broader framework wherein uncertainty measurement is understood as a decision.

---

[33] Indeed, rarely is one concerned with the abstract instance of selecting an individual in some unspecified way. In most cases, there is circumstantial information that influences one's assessment of the probability that the person of interest is drunk, different from what the naked frequency data may suggest. For example, the person may exhibit suspicious driving behavior (such as speeding, unusually slow or insecure driving), have a vehicle showing recent damages or a missing number plate, etc. Interestingly, Dahlman himself considers the use of a base rate problematic in this example ("It may not be very realistic to imagine a situation where 60% are driving under the influence (...)", supra note 13, at 21).

[34] Dennis V. Lindley, Understanding Uncertainty (2006), at 103. The same point has also been made by De Finetti (2008), supra note 25, at 128: "The frequency with which certain events obtained or will obtain cannot be identified with probability. Frequency is a mere fact, independent of both the meaning of probability and the probability values assigned to the events."

[35] A formal argument to show that relative frequencies, where available, can be coherently related to a person's probability for an event of interest relies on the notion of exchangeability. The technical details are beyond the scope of this paper. For a discussion see, e.g., Franco Taroni et al., Reconciliation of Subjective Probabilities and Frequencies in Forensic Science, 17 Law, Probability & Risk 243–262 (2018).

B. *Likelihood and probability are not synonyms*

Legal scholarship on the use of probability in the law is further complicated by the confusion between likelihood and probability. Indeed, they are often treated as synonyms, or near synonyms. In everyday language, but also in the discourses of many learned scientists, saying that an outcome or event is 'likely' is often taken to mean that it is 'probable'. Dahlman, for example, mentions "estimated likelihoods of the evidence"[36] and, more specifically, "the likelihood of a false positive, i.e. the probability that such a match would occur with (...) a randomly picked innocent person."[37]

The difference between likelihood and probability is technically subtle, and conceptually important. From a formal point of view, the *probability* Pr of an event or outcome E conditional on hypothesis H, written $\Pr(E|H)$, is also referred to as the *likelihood* of H, given E, sometimes written $L(H;E)$. The term likelihood relates to the hypothesis, though, not the evidence, which is why the above expressions 'likelihood of the evidence' are ambiguous. To further illustrate the distinction, let E denote the event of finding large and fresh bloodstains (with a DNA profile corresponding to that of the victim) on the jacket of a person of interest (POI), and H the proposition that the POI assaulted the victim. It appears fair to assert that the probability of observing E, the POI's jacket showing fresh bloodstains with a DNA profile corresponding to the victim, conditional on the POI being the person who assaulted the victim (*H),* is high. Let it be close to 1 for the sake of illustration, which is reasonable when the POI is arrested shortly after the assault. Thus, the likelihood of H given E, i.e. $L(H;E)$, is high: the proposition H provides a good account of the findings E. But that does not mean that $\Pr(H|E)$, the probability of H (the POI being the offender) conditional on E is high, too. Only $\Pr(E|H)$ is. To help understand why this is so it is useful to consider that there may be other viable alternative propositions that also have a high likelihood. Consider, for example, the alternative proposition H' according to which the POI is not the offender, but is a good Samaritan who provided first aid to the victim. Under this conditional, the probability of the bloodstain evidence E may be high as well. That is, $\Pr(E|H')$ may also be close to 1. But again, this high likelihood of H' does not imply that H' is (highly) probable. Indeed, in the example here, the likelihoods of H and H', given E, are both high (i.e., close to 1), i.e. they both provide good accounts of the findings E, but clearly H and H' cannot be both highly *probable* (i.e., close to 1) for as this would lead to a total probability greater than 1, and hence violate the rules of probability. Stated another way, with likelihoods, the evidence is fixed, but the propositions vary. Using again the above example, the two

---

[36] Supra note 13, at 15.
[37] Supra note 13, at 20.

conditional probabilities $\Pr(E|H)$ and $\Pr(E|H')$ feature the same evidence E under different conditionals, H and H', and are two likelihoods. This is commonly known among forensic scientists who refer to the ratio of the two conditional probabilities $\Pr(E|H)$ and $\Pr(E|H')$ as the likelihood ratio.[38] Inherent in this notion is that the term likelihood refers to a proposition H with respect to E, but not the reverse: i.e., there is *no* reference to a 'likelihood of the evidence E'.

More generally, the distinction between the likelihood of the proposition given evidence, and the probability of the same proposition conditional on the same evidence is important to raise awareness about what is known as the "fallacy of the transposed conditional",[39] described also in legal contexts.[40] A fallacious transposing of the conditional occurs, for example, when one equates the probability of observing a positive reaction to a diagnostic test if a person handled a certain substance (e.g. explosives) and the probability of the person having handled said substance if a positive reaction is observed.[41]

Our point here is not to claim that the distinction between probability and likelihood is a safeguard against fallacious conclusions in general. Indeed, as argued by Sullivan, "[t]he difference between these concepts is too foreign and too esoteric for mere words to make any difference on their own."[42] Notwithstanding, it is a general hindrance to both legal research and practice if the underlying concepts are not properly distinguished.

C. Flat earth and moon landing: the myth of 'true' probabilities

There is one more salient topic in the expression "estimated likelihoods of the evidence"[43]. It has to do with the use of the terms 'estimate' and

---

[38] E.g., Colin Aitken & Franco Taroni, Statistics and the Evaluation of Evidence for Forensic Scientists (2004). Likelihood ratios have a long history in statistical literature where they are known as a measure of the value of experimental findings (see, e.g., Irving J. Good, Probability and the Weighing of Evidence (1950) at 63).

[39] Persi Diaconis & David Freedman, The Persistence of Cognitive Illusions, 4 Behavioural and Brain Sciences 333–334 (1981) at 333.

[40] William C. Thompson & Edward L. Schumann, Interpretation of Statistical Evidence in Criminal Trials: The Prosecutor's Fallacy and the Defense Attorney's Fallacy, 11 Law and Human Behavior 167–187 (1987).

[41] For a discussion of this fallacy in the context of the Birmingham Six bombing case, see Bernard Robertson & Tony Vignaux, Interpreting Evidence, Evaluating Forensic Science in the Courtroom (1995) at 92. For an example in medical diagnosis see e.g. David M. Eddy, Probabilistic Reasoning in Clinical Evidence, Problems and Opportunities, in: Daniel Kahneman, Paul Slovic & Amos Tversky (Eds.), Judgment under Uncertainty, Heuristics and Biases, 249–267 (1982) at 254. General guidance is also provided by The Inns of Court College of Advocacy & The Royal Statistical Society, Statistics and Probability for Advocates: Understanding Statistical Evidence in Courts and Tribunals (2017).

[42] Sean P. Sullivan, A Likelihood Story: The Theory of Legal Fact-Finding, 90 University of Colorado Law Review 1–66 (2019).

[43] Supra note 15, and similarly in Fenton et al., supra note 26 ("estimating what we call

'estimation', commonly used in sentences such as "a (...) probability of the guilt-hypothesis can be estimated".[44] Indeed, it is common to encounter expressions like 'estimating a probability' in everyday language, but again, complications are in the details, and their conceptual implications are far reaching. In this Section, we will point out why probabilities are not the object of 'estimation'.

Start by considering a situation in which the term 'estimation' *is* useful and appropriate: it is the case when there is a real quantity to discover. Take the example of the proportion of members in a target population that have a particular feature, such as a eye colour. In theory, it is possible to determine such a proportion by inspecting every member of the population, though this is time and resource consuming.[45] It is for this reason that scientists rely on the inspection of only a selected number of members of the population, in order to draw a conclusion about the population as a whole – in terms of an estimate. This bears no relationship, however, with a proposition such as 'the defendant is the offender'.[46] This proposition is either true or false. Here, probability merely serves as an expression for your (or, anybody's) uncertainty that this proposition is true. This probability can take any value between zero and one, but there is *no* 'true' or 'uniquely/objectively right' probability, i.e. your belief, that could somehow be 'estimated'. All that can be said is that the more you are sure that a given proposition is true, the higher your probability (i.e., close to one), and, conversely, the surer you are that the proposition is false, the smaller your probability (i.e., close to zero). Probability is a very liberal concept[47] in that it does not tell one what their probability should be (besides that it ought to be a number between zero and one). This property raises many concerns among legal and forensic science scholars that deserve careful consideration.

A main concern is that personally-assigned probabilities – commonly referred to as subjective probabilities – are prone to "abuse (...) to allow people to render opinions (...) on what is sometimes no better than a guess."[48]

---

the 'opportunity prior probability'; the probability that a suspect is guilty", at 2).

[44] Supra note 13, at 16.

[45] Also, the population would change during the course of inspection through births, deaths and migration (Ian W. Evett & Bruce S. Weir, Interpreting DNA Evidence (1998) at 43).

[46] Note that we avoid the term 'guilt-hypothesis' because, as noted above (supra note 28), guilt is a decision, not a hypothesis.

[47] Dennis V. Lindley, Making Decisions, Second Edition (1985) at 22.

[48] D. Michael Risinger, Reservations About Likelihood Ratios (and Some Other Aspects of Forensic 'Bayesianism'), 12 Law, Probability & Risk 63–73 (2013) at 71. The term 'guess' is also used by Dahlman, supra note 13, who notes that "[d]ecision-makers are on their own, making guesses about reference classes" (at 20).

Similarly, Allen is concerned about "uncontrolled subjectivity"[49] and that probabilities "could be any numbers at all, and they need not be constrained in any way by the quality of the evidence".[50] Forensic scientists also show resistance against assessments "which are directly the result of subjective judgement".[51] These are valid concerns with respect to the liberal concept of probability and they merit careful attention.

We have elsewhere critically exposed what we termed "*unconstrained subjectivism*"[52], i.e. an arbitrary and speculative use of probability to support "fanciful probability statements devoid of any meaningful justification".[53] This, however, is not how serious personal probabilists understand probability.[54] In this context, subjective does not mean arbitrary. Subjective probabilists invoke particular criteria to constrict arbitrary probability assignment. Specifically, there are two related, critically important notions – justification and conditioning information – that need to be considered in personal probabilities.[55]

To illustrate, consider two examples. The first deals with the proposition that the earth is flat.[56] Suppose that you assign a high probability for this proposition to be true. This may seem like a silly example and, at first sight,

---

[49] Supra note 6, at 138.

[50] Supra note 5, at 10. Note, however, that even the purest subjective probabilists are not dismissive of evidence. For example, de Finetti, supra note 7, requires that our beliefs have a "reasonable relationship to any kind of relevant objective data" (at xv).

[51] Geoffrey S. Morrison & Ewald Enzinger, What Should a Forensic Practitioner's Likelihood Ratio Be?, 56 Science & Justice 374–379 (2016) at 374.

[52] Alex Biedermann, Silvia Bozza, Franco Taroni & Colin Aitken, The Meaning of Justified Subjectivism and its Role in the Reconciliation of Recent Disagreements Over Forensic Probabilism, 57 Science & Justice 477–483 (2017) at 478.

[53] Id.

[54] As noted by Bruno de Finetti, Probability, Induction and Statistics (1972): "We strive to make such judgments as dispassionate, reflective and wise as possible (...)" (at 144). According to de Finetti, every probability elicitation comes down – in essence – to an individualized assessment, thus arguing in favor of "removing the distinction between certainty and guessing, which have been identified as fallacious notions" (at 62), and opposing to "(...) the artificial distinction between knowing and guessing as distinct categories of response" (at 63).

[55] See, e.g., Paolo Garbolino, Explaining Relevance, 22 Cardozo Law Review 1503–1521 (2001), invoking the notion of "[r]easonable personal probabilities" (at 1504), and arguing that "[i]f we cannot require everybody sharing the same likelihoods, we can require everybody having *justified* likelihoods" (at 1506). Interestingly, Swiss criminal law, for example, is more demanding: it mandates the evaluation of evidence to be rational, which in turn is defined by some legal scholars as being 'intersubjective', in the sense that multiple judges or courts could agree on the assessment of the evidence that was done in a given case. See Jean-Marc Verniory, La libre appréciation de la preuve pénale et ses limites, 118 Revue Pénale Suisse (2000), 378–413, at 390–393.

[56] This example is adapted from Lindley (2006), supra note 34, at 13, but we develop it further here in order to illustrate more points of discussion.

could serve as perfect illustration that probability indeed supports fanciful statements. But, as Lindley notes, "there is nothing in our rules [of probability] to say you are wrong, merely that you are unusual (...)".[57] This leads to several points of reflection. First, we should ask why the assertion of a high probability for the proposition that the earth is flat seems inappropriate to us. Presumably, it seems inappropriate to us because of the overwhelming evidence we have to the contrary.[58] So, whenever we encounter a person who we think holds strange or unusual beliefs, we should question that person on her evidence. It may well be that the knowledge base of that person does not contain the same evidence as that available to us, that the person of interest may assess available evidence differently, or consider evidence only selectively.[59] But consider now a variation to this example and imagine a person back in ancient times, before the rise of Hellenistic astronomy. Given the evidence available at that time, it would not appear unusual to the casual observer to believe that the earth is flat: after all, her observations are limited to the perspective of a person standing on the earth with her two feet. When we judge the appropriateness of such a person's high probability for the event that the earth is flat, we need to keep in mind the conditioning information: today's body of knowledge is different from the one available in ancient times. Whatever the person's views are on the proposition of interest, probability is not at fault in any way – rather, it works just fine: it serves as a precise expression of what a given person thinks about the truth of a proposition of interest. And, in all cases – even when the assertions and beliefs seem obviously right – it is worthwhile to inquire about the foundations of probability assertions, i.e. how they are justified, and on which conditioning information they rely.[60]

As a second example, consider the possibility of a manned mission to the moon in the near future. What is the probability of the successful arrival of a manned spacecraft on the surface of the moon?[61] Before addressing this question, we might wish to obtain some additional information, such as who

---

[57] Supra note 34, at 13.

[58] E.g., the website https://www.nasa.gov/topics/earth/images/index.html hosts images provided by the NASA.

[59] For example, the person may distrust public sources and governmental agencies, such as NASA.

[60] As an aside, note that the flat earth example also shows that probability assignments cannot be substantiated with base rate or frequency perspectives: we cannot meaningfully conceive of a series of earths and count the number of times the earth was found to be flat and spherical, respectively.

[61] As in the previous example, we insist on the fact that there is no base rate to consider here. We have to deal with a unique singular, non-repeatable event that is different from previous space missions in various respects, such as the features of the spacecraft, the crew and the high number of other factors that characterise modern space missions.

is planning the mission. Is it an experienced and well-funded entity, such as NASA, or a small country, such as Liechtenstein, that has no experience in space exploration? Or, is it a private company? This example shows that we naturally inquire about the relevant conditioning information when assessing a probability, and that we care about this information to be thorough and trustworthy, in order for our probability assignment to be justifiable. But again, there is no single 'right' probability: the information available to an engineer involved in said space mission might well be different from the information available to the general public. It will not be surprising, thus, that the probabilities assigned to the proposition by engineers or members of the public will differ: while each opinion may be well justified in its own right, the conditioning information between different people will vary.

However, it may rightly be invoked that probability does not prevent a person from declaring a more extreme probability than what this person actually considers to be a suitable probability in the case at hand.[62,63] It is for this reason that the notions of justification and conditioning information are powerful notions to help examine whether probability assessors are in possession of information that entitles them to their assertions, in law[64] as any other context of application.

D. The veil of robustness and endless regress

Related to the topic of justification of probability assertions are concerns about the extent of supporting information. Dahlman notes: "A further problem with probability assessments that are based on insufficient reference data is that such assessments are 'un-robust'. That an assessment is un-robust means that it is likely that additional reference data would change the assessment significantly, and this is not acceptable."[65] These assertions confuse different issues.

First, while available information may vary to any extent, probability incorporates as much relevant information as is available. But, if there is limited information, this is not a fault or problem of probability, but merely reflects an intrinsic condition of the real world.[66] Even if one does not endorse

---

[62] In later parts of this paper we will expose how the decisional conceptualization of probability specifically copes with this challenge.

[63] In forensic science, for example, such overstatements became known as "leap of faith" (David A. Stoney, What Made us Ever Think We Could Individualize Using Statistics? 31 Journal of the Forensic Science Society 197–1999 (1991) at 198).

[64] As noted by Allen (2017), supra note 6, at 137, "the evidentiary process requires personal knowledge or justified expert knowledge."

[65] Supra note 13, at 20.

[66] Conversely, new evidence is also an inherent feature of the real world and is the reason why modern legal orders permit appeals, with some systems permitting secondary and subsequent appeals.

probabilistic thinking, the issue of imperfect data will persist. The notion of justified probability assertion is useful in this context because it is transparent regarding its foundations and, thereby, informs discussants about the extent to which probabilities rely on data. In any case, decision makers preserve their liberty regarding how to act or proceed based on such probabilities.

Second, probability assessment is bound to systemic conditions, i.e. the properties of the context of application. As Allen insists, "[t]he law has to decide at the moment in question; it cannot suspend belief while a series of tests is conducted to see if anything useful emerges."[67] This leaves decision makers with no choice other than to make the best use of whatever relevant information is available at the time a decision needs to be made. As noted previously, probability is not a magic device that can 'produce' information out of nothing, it merely serves the purpose of expressing a person's uncertainty as a function of available data, however good or poor that data may be.

Third, invoking other data that might be observed, but that are not, is a typical concern of frequentist statistics that postulates that observations could conceivably have resulted in another set of data.[68] Underlying this concern is what Lindley[69] calls hindsight, which raises the question of how to look at *ex ante* assessments from an *ex post* data perspective. For example, it is commonly said that when new data suggest or favor propositions other than those endorsed ex ante, the previous opinion (or, probability) is incorrect or wrong. However, such critiques are misplaced because, by definition, a probability is conditioned on the information available at the time when the original assessment was made. It is pointless to criticize an assessment based on information that was unavailable to the probability assessor ex ante. The point of subsequently obtained data is not one of judging probabilities as being 'right' and 'wrong', or making a correction, but one of recognizing that a former probability assessment needs readjustment in the light of new information, leading to a new probability assignment.[70,71]

----

[67] Allen (2017), supra note 6, at 139, and Allen & Stein (2013), supra note 3, at 601.

[68] We have noted earlier in § A. why we do not consider events of legal interest to be conceivable in terms of repetitions.

[69] Supra note 47, at 192.

[70] See also De Finetti (2017), supra note 7, at 176: "If, on the basis of observations (...) one formulates new and different previsions (...) for events whose outcome is unknown, it is *not* a question of a *correction*. (...) new results (...) enrich one's state of information, drawing out of this the evaluations *corresponding to this new state of information*. For the person making them (You, me, some other individual), these evaluations are as correct *now*, as were, and are, the preceding ones, though of *then*." (emphasis as in original)

[71] As an aside, note that this view is also in agreement with foundational principles in statistical theory according to which inferences ought to rely exclusively on actually observed data, not other data that might have been observed but were not (e.g., James O. Berger & Robert L. Wolpert, The Likelihood Principle (1984)).

An outgrowth of concerns about the extent to which probabilities are supported by data are proposals of interval or second-order[72] probabilities. In a nutshell, these proposals seek to circumvent the difficulty of asserting a probability – as an expression of one's uncertainty – in terms of a single number. Recently, this has led to considerable controversy among forensic scientists, with some discussants insisting on the logic of understanding probability as a single number,[73] and others arguing to the contrary.[74] While, as per definition, probability is a *single* number – for different numbers express *different* states of uncertainty – the practical implementation of this definition at its full level of rigor may be difficult. However, it is not a viable solution then to merely add an additional level of technicality in terms of intervals or the like. Indeed, if one is not able to assign a single-valued probability in the first place, how is one to assign a probability, or an interval (requiring *two* values, an upper and a lower bound), *over* a probability? According to Lindley, this leads to an infinite regress[75] and confuses the measuring device (i.e., probability) with its application in practice.[76]

E. *The struggle with the '(non-) knowledge' of probabilities*

The concerns raised in the previous sections regarding ostensibly true and robust probabilities reveal a further question that is often raised in connection with the assertion of probabilities, in particular when they are hard to elicit: the issue of how one can 'know' probabilities, especially knowing which is the 'right' probability. Indeed, one often hears statements such as 'I don't

---

[72] See, e.g., supra note 13, at 20: "Robustness can be modelled as a second-order probability (...)".

[73] E.g., Franco Taroni et al., Dismissal of the Illusion of Uncertainty in the Assessment of a Likelihood Ratio, 15 Law, Probability & Risk 1–16 (2016), Anders Nordgaard, Comment on 'Dismissal of the Illusion of Uncertainty in the Assessment of a Likelihood Ratio' by Taroni F., Bozza S., Biedermann A., Aitken C., 15 Law, Probability & Risk 17–22 (2016), Charles Berger & Kas Slooten, The LR Does not Exist, 56 Science & Justice 388–391 (2016).

[74] For a collection of forensic science articles on the topic see e.g. Geoffrey S. Morrison (Ed.), Measuring and Reporting the Precision of Forensic Likelihood Ratios, Virtual Special Issue, Science & Justice, www.sciencedirect.com/journal/science-and-justice/special-issue/102F0FGVD03. The practical impact of these discussions appears to remain limited, however, as forensic scientists continue to make debatable assertions about intervals (for a recent example see, e.g., Dennis McNevin, Kirsty Wright, Janet Chaseling, Mark Barash, Commentary on: Bright et al. (2018) Internal validation of STRmixTM – A multi laboratory response to PCAST, Forensic Science International: Genetics, 34: 11–24, 41 Forensic Science International: Genetics, e14–e17 (2019)).

[75] "(...) it is nonsense for you to have a belief about your belief if only because to do so leads to an infinite regress of beliefs about beliefs about beliefs" (supra note 34, at 115). The notion of 'infinite regression' has previously been used by Bruno de Finetti, supra note 54, at 193.

[76] Supra note 34, at 36.

know what *the* probability is', where the emphasis on 'the' suggests that there is a real and uniquely applicable – and hence potentially *knowable* – probability to be discovered. This may occur, for example, when one attempts to define and assess probabilities in terms of base rates, but one cannot provide or agree on the applicable reference classes, or one lacks suitable base rate data. But again, this confuses the assessment of probability with its definition. If one can agree on the view that probability is, in a most general sense, an observer's expression of the degree of belief that a particular uncertain event is true, depending on the specific information and knowledge available to that observer, then assertions of non-knowledge of probabilities turn out to be a "pseudo-problem".[77] According to this account, probabilities are an observer's justified position statements regarding events of any kind, hence asserting that one does not know one's probabilities would mean to say that one does not know one's states of mind (i.e., opinions):

> "Among the answers that do not make sense, and cannot be admitted, are the following: 'I do not know,' 'I am ignorant of what the probability is,' 'in my opinion the probability does not exist.' Probability (or prevision) is not something which in itself can be known or not known: It exists in that it serves to express, in a precise fashion, for each individual, his choice in his given state of ignorance."[78]

The challenge, thus, does not rest with the concept of probability, but with limitations of one's commitment to the task of probability elicitation, or with external factors that prevent one from asserting probabilities. These factors are multiple and possibly interrelated, such as aversion against publicly exposing one's (probabilistic) view, and bearing the responsibility that comes with probability assertion. There may also be policy reasons or legal obligations that prevent one from the reliance on and exposure of relevant information, for example when it is proprietary.

F. The trap of murky propositions

Questionable elicitation targets, in the form of highly peculiar hypotheticals, are a further source of difficulties in assessing probabilities, and are often used in attempts to illustrate the limited suitability of personal probabilities. Vermeule, for instance, invoking an example drawn from Elster[79], argues that questions such as whether "Norway in the year 3000 will

---

[77] De Finetti, supra note 54, at 63. At another instance, De Finetti, supra note 7, has noted: "Probability is the result of an evaluation; it has no meaning until the evaluation has been made and, from then on, it is known to the one who has made it. For this obvious reason alone, the phrase 'unknown probabilities' is already intrinsically improper (...)" (at 122).

[78] Supra note 7, at 72.

[79] Jon Elster, Explaining Technical Change (1983) at 199.

be a democracy rather than a dictatorship"[80] lead to meaningless probabilities, devoid of any epistemic credentials.

It is true that there is nothing in probability theory that prevents one from giving a probability in reply to such a question,[81] and it may seem appalling that personal probabilists consider it feasible – in principle (though not necessarily straightforward) – to give a probability here. However, asserting a probability, even for strange propositions, does not found a criticism of probability. Sensible subjective probabilists might well invoke that they do not wish to declare a probability because of the nature of the question asked and the intricacy of the domain of expertise on which they are consulted – rendering *justified* probability assignments for the case at hand too challenging.

Thus, the difficulty relies in the real-world problem under study, not with the chosen analytical framework of thinking. As noted by Friedman, probability "(...) is a flexible template. It can take into account as much complexity as its user is able to handle",[82] and, to which we can add, that the user is willing and feels able to defend. What is more, Friedman argues,[83] the theory would indeed have limited value, and lack comprehensiveness, if it did not – in principle – allow us to contemplate any real-world proposition, however peculiar and difficult to assess it may be. In this sense, there is nothing inherently wrong with asserting a personal probability for the 'Norway proposition' as long as all the discussants are clear about the idiosyncrasy of the elicitation target (i.e., an event in the far future) and the limitations in the available knowledge and expertise required to appropriately address the question under scrutiny. As an aside, it is also clear that resorting to other types of probability, such as frequentist or base rate inspired interpretations, would not be an option. Among the remaining alternatives would be not to opine on the question of interest, though we have mentioned previously that in legal contexts, fact-finders do not have the liberty to delay action.

G. *Intermediate summary: from abstract definitions to probability as a decision*

The above sections have raised and discussed a number of recurrent, but by no means exhaustive, criticisms of probability. The objections we have reviewed appear to be influenced by a frequentist definition of probability. We have seen that attempts to apply this definition of probability runs into

---

[80] Adrian Vermeule, Judging Under Uncertainty, An Institutional Theory of Legal Interpretation (2006) at 173.

[81] A similar comment has been made previously in §1.C. about the flat earth example.

[82] Supra note 24, at 288.

[83] Id.

operational difficulties. These range from the definition and availability of relevant data for a case at hand to conflicts with principles and values of the legal process. Specifically, we have identified as one source of difficulty the fact that classic (i.e., frequentist) pursuits of probability conflate the definition of probability with its assignment. This has led us to provisionally favour the counter-position according to which probability is a person's individual assessment of uncertainty regarding any event, conditioned on whatever data and knowledge is available to this person at the time uncertainty is (probabilistically) quantified.

This interpretation of probability as belief is neither a new or an original standpoint, nor is it free from attracting its own objections. Among the chief objections is the concern that the liberal concept of probability which, in a strict sense, places only limited constraints on probability assertions (i.e., coherence requirements), can induce people to assert any probabilities they wish, without any epistemic warrants. In §I.C. we argued that such unconstrained probabilism, albeit a valid concern, does not correspond to how serious probabilists understand personal probability. They consider that the liberal concept of probability comes at the price of responsibility that must be assumed.

In the remaining parts of this paper, we shall illustrate how an individual's sense for responsible probability assertion can formally be incentivised. We will do so by relying on a particular concept, to be explained in due course, which mirrors probability through the way in which a given individual is prepared to act under uncertainty, given their best knowledge. The rationale for this methodological choice derives from an analogy with the burden of decision at other instances of the legal process. Conclusions regarding ultimate issues, for example, are – by their nature – questions of *decision* which ought to be argumentatively well grounded. Similarly, we will argue, one can interpret the sensible assertion of a probability regarding an intermediate uncertain proposition as a decision. That is, an individual *decides* which probability to assert in any given instance, by making the best use of the entirety of the available knowledge and information, and by aiming to appraise reality as defensibly as possible.

PART 2: THE DECISIONALIZATION OF PROBABILITY

A. The possibility of measuring uncertainty using a standard

Unlike accounts that seek to conceptualise probability in terms of abstract and impersonal notions (e.g., secondary concepts such as reference classes, base rates, etc.), we shall start the presentation of the decisional perspective of probability by placing the individual in the centre of our analysis. This does not deny the view that factfinders ought to be disinterested individuals,

and the court be "a disinterested forum",[84] instructed to pursue legal proceedings as "the disinterested application of common sense"[85] including "all the necessary forms–deductive, inductive, abductive statistical–as necessary or appropriate".[86] The point of interest here is the uncontested fact that all legal trials place decision-making at their centre, an activity that is fundamentally directed by individuals, mandated to implement societal values. The choice of an analytical framework should thus properly account for this reality. The person-centric approach precisely characterises this view "(...) for it describes a real situation of a subject, or observer, contemplating a world, rather than talking about a world divorced from the observers living in it, as do some sciences."[87]

However, before we proceed with exposing the decisional perspective to probability, it is necessary to address one more intricacy. It deals with the question of whether, and if so, how, appraisals of uncertainty can be given a clear, firm and *measurable* structure – an aim that the proclaimed objectivistic approaches based on single notions, such as base rates, attempt to do. Because, for a concept to be tractable and scientifically sound, one should be able to measure it, using numbers – at least conceptually.[88]

The property of measurability is based on the feasibility of comparisons, a notion that is commonly known and recognised by many cognitive concepts, beyond probability. For example, the relative plausibility account involves qualitative assessments such as "(...) the common sense judgement that some explanations are on occasion considerably better, not just better, than others"[89]. Similarly, probabilists make comparative statements of the kind 'I am less sure that event A holds rather than event B', where A and B are propositions about which the individual has incomplete knowledge, and where the degree of belief in the truth of each event may be expressed by a number between 0 and 1, i.e. the person's probabilities. But how do we get to the understanding that the admissible numbers for probabilities lie in the range between 0 and 1, and how do we get to consider these numbers as our

---

[84] Ronald J. Allen, Burdens of Proof, 13 Law, Probability and Risk 195–219 (2014) at 196.

[85] Ronald J. Allen, A Note to My Philosophical Friends About Expertise and Legal Systems, 28 Humana.Mente Jounral of Philosophical Studies 71–86 (2015) at 79.

[86] Id. at 79.

[87] Lindley, supra note 47, at 20.

[88] Lindley, supra note 47, for example, favors a numerical approach to uncertainty because "(...) number is the essence of the scientific method and it is by measuring things that we know them" (at 13). The idea of measurability has parallels to the notion of *operational definitions* developed in other fields, such as physics (see, e.g., Percy Bridgman, The Logic of Modern Physics, 1927, and Frank Lad, Operational Statistical Methods: A Mathematical, Philosophical, and Historical Introduction, 1996).

[89] Supra note 84, at 218.

expression of uncertainty? There are various ways to arrive at this result, but we shall mention here only one way, mainly because it uses only minimal assumptions and technicalities. This should minimize the potential for objections and make the presentation accessible for a general audience. The argument is due to Lindley[90] and relies on our ability to compare uncertain events, as mentioned above. But rather than just comparing any two events, and assert which of the two compared events is endowed with more or less uncertainty than the other, Lindley introduces the idea of a comparison with a standard. That is, in order to measure our uncertainty regarding a given event (the measurand) numerically, a comparison is to be made against a standard. The standard is set up carefully such that its properties should be agreed by every reasonable person.[91] The conceptual standard introduced by Lindley is a container with a proportion p of black balls, and the 'standard' event is the observation of a black ball after a single draw.[92] Clearly, if all the balls in the container are black (and hence the proportion p is 1), we will be sure to observe a black ball: there will be no uncertainty about this observation. Similarly, if the proportion p of black balls is zero, we will be certain not to observe a black ball after a single draw. It is now obvious that our uncertainty regarding the ball being drawn to be black is closely related to the proportion p of black balls in the container, which can be adjusted to any number between 0 and 1. Note also that the standard can be defined in a more or less finely graded way, depending on the total number of balls: a container with 100 balls allows uncertainty to be expressed in 1% increments, whereas an urn with 1000 balls allows uncertainty to be expressed in increments of 0.01%, and so on.

We can now use this standard for measuring our uncertainty about any event of interest by asking ourselves how our uncertainty about the uncertain event, the measurand, *compares* to the uncertainty of drawing a black ball from the container – our standard.[93] Clearly, the answer depends on the proportion p of black balls. In particular, there must be a proportion p such

---

[90] Supra note 47, 17–20.

[91] Note that the idea underlying this procedure is analogous to the one used for the measurement of physical quantities, that is by comparing the measurand against some standard unit.

[92] Note that the ball is only to be drawn once, and the observation of a black ball is solely determined by the proportion p of black balls. There is no notion of relative frequency involved here. As noted forcefully by Lindley, supra note 34, the "(...) long-run frequency (...) in repeated drawings is irrelevant. After its withdrawal, the urn and its contents can go up in smoke for all that it matters" (at 36).

[93] Our presentation here is simplified in the sense that it does not mention Lindley's device of considering the comparison in terms of a choice between two options (with not stake involved): receiving a small prize if the target event is true and to receive the same prize if the ball from the container turns out to be black.

that our uncertainty about the target event to be measured is the same as our uncertainty regarding the drawing of a black ball from the container. Once determined, this point of equivalence, as given by the proportion *p*, is then said to be our probability for the uncertain event.[94]

To be clear, the suggestion here is not that we carry an urn with us to conduct a probability elicitation whenever we need one.[95,96] The comparison is only to be contemplated. The sole purpose is to provide a way to explain, in principle, the nature and the form of probability assessments, using only most basic ingredients – free of any difficulties arising from mainstream (relative) frequency ideas. We do not deny that – despite its simplicity – the procedure of comparison with a standard may be difficult to apply, even conceptually. Indeed, critical observers might say that there is nothing in the procedure to prevent aleatory and fanciful assertions of equivalence points and, worse, that there is no guidance to help one find out *when exactly* one has reached a point of equivalence, however well-intentioned and committed one may be in the application of the procedure.

These objections have to do with unconstrained subjectivism, a problem that we addressed in §I.C. Probability being a liberal concept, it requires scrutiny and challenge in its application. We will provide further means to address this issue in the next section. The objection regarding the difficulty of measuring uncertainty probabilistically is a valid one, though it is important to be precise about where exactly the difficulty lies. As argued in §I.F. the difficulty is not due to the measurement procedure, which is an extremely simple one, but to the complexity of the measurand, that is the real-world event under consideration.

A further objection that may be raised concerns the relationship with evidence. As the attentive reader may have observed, the above definition of probability does not explicitly mention evidence and data. It is tempting, thus,

---

[94] The balls in the urn should not be misunderstood as a representation of a collection of real-world cases. The urn's sole purpose is to serve as a device for conceptualizing an event with a well-defined probability; it is this uncertain event which is then to be *compared* to the real-world event (or, proposition) of interest – the measurand – but which could be any, and has no relationship with a ball in an urn.

[95] The measurement procedure using a standard is not entirely theoretical, however, as some researchers have successfully used it to elicit probabilities in experiments with prospective jurors (see, e.g., Nicholas Scurich, Kenneth D. Nguyen & Richard S. John, Quantifying the Presumption of Innocence, 15 Law, Probability & Risk 71–86 (2016)).

[96] Interestingly, practicing lawyers facing strategic decisions more openly embrace conceptual devices for probability elicitation. For a method comparable to the urn scheme, but based on a betting (or, probability) wheel with colored segments of varying size, reflecting the 'chance of winning', see e.g. Howell E. Jackson, Louis Kaplow, Steven M. Shavell, W. Kip Viscusi & David Cope, Analytical Methods for Lawyers (2003), at 28. More detailed expositions can also be found, e.g., in Detlof von Winterfeldt & Ward Edwards, Decision Analysis and Behavioral Research (1986), at 117.

to interpret this as yet another argument in support of the view that personal probabilities lack proper evidentiary bases. However, this reproach is misplaced because, as we have clarified at the outset, we are making a distinction between the definition of probability and its evaluation. The above definition – in terms of the conceptual application of a standard – only serves to express the idea of measuring uncertainty through probability. It is assumed that personal probability assignments thus made will take into account any relevant data there may be. This, however, is a question of how to evaluate probabilities in particular instances, and there are formal procedures available for this task.[97]

It is legitimate to ask whether these are satisfactory answers that make subjective probability any better and more viable than classic (e.g., frequentist) accounts. Specifically, a key concern that remains – relevant for any conceptualization of probability – is the question of how 'good' a given probability assignment is, and how one can give incentives so that probability assessors give undistorted, non-exaggerated and non-aleatory probability assignments. The decisionalist perspective to probability, presented in the next section, precisely addresses this point.

B. *Understanding probability assertions as decisions*

Inherent in the understanding of probability as a person's expression of their degree of belief that a certain proposition is true, is the avoidance of deference to abstract notions (e.g., relative frequencies and base rates) that attempt to attribute to probabilities any sort of existence independent of the person who evaluates them. As we have seen, many scholars in law and other fields interested in probability have repeatedly exposed problems arising from said notions. While, technically, we agree with the impasses they expose, we disagree with their conclusion that probability, especially personal probability, is an inadequate concept; such a conclusion is tantamount to throw up one's hands at the price of making assessments at all. As we see them, the criticisms arise from unfruitful attempts of using single notions for both, the definition of probabilities in general, *and* their evaluation in actual instances of application. The interpretation of probability as measures of degrees of belief, as exposed in the previous section, effectively avoids such complications.[98] The question we address now is, given the understanding of probability being defined as a person's degree of belief, how can we conceive of a way to encourage individuals to make probability assignments in a serious and well reflected way.

One possibility consists in appealing to a person's responsibility for their

---

[97] See also supra note 35.

[98] Maria C. Galavotti, Who is Afraid of Subjective Probability? in: C. de Florio & A. Giordani (Eds.), From Arithmetic to Metaphysics (2018), 151–157.

probability assignments by insisting on the view that the assertion of a probability actually amounts to making a decision – a decision that ought to be 'well done', in some sense. To clarify this idea, the probabilist Bruno de Finetti emphasized the idea of scoring: the person who indicates her probabilities shall do so by keeping in mind that she will incur a penalty (or, score) based on the number (i.e., probability) that she has reported. The scoring method – in literature known as a scoring rule – is designed in such a way that it is in the person's interest to state the probability that corresponds to her actual degree of belief. Stated otherwise, under this kind of scoring rule,[99] the person can, by following the precept of providing sincere probability assertions, minimize the expectation of the penalty. In a nutshell, the score – we choose here the quadratic score as an example[100] – works as follows: for an event E, which may be either true or false, represented by the numbers 1 and 0, respectively, the asserted probability α for event E receives the score $(1-\alpha)^2$ if the event E turns out to be true, and α^2 if the event E turns out to be false. The score thus amounts to the squared difference[101] between the asserted value α and the truth-values 1 and 0.

It might be objected at this point that the design of the scoring rule is of limited use for applications in the law because ground truths are typically unavailable in this context. However, as will be explained later in this section, actual ground truths are *not* needed, only (personal) probabilities for states of nature. These will be used to compute so-called probability-weighted scores, which in turn represent the core of the scoring rule argument.

Before approaching these details, start by considering the following more general properties of this scoring regime and their implications for probability utterances:

- Broadly speaking, we would say that one's thought, expressed in terms of probability, is 'good' if it lies in the direction of truth.[102] As may be seen, the quadratic score leads to smaller penalties the closer one's asserted probability is to 1 and 0 for events that are, respectively, true and

---

[99] Rules with this property are called proper scoring rules. Other scoring rules exist, but they may not guarantee that stating one's probability sincerely is the optimal decision (i.e., will minimize the expected score).

[100] For other examples of proper scoring rules, such as the logarithmic scoring rule, see e.g. Giovanni Parmigiani & Lurdes Inoue, Decision Theory, Principles and Approaches (2009) at 195.

[101] Note that taking the simple difference, rather than the squared difference, would not be a proper scoring rule (i.e., under such a rule, honest reporting of probabilities does not minimize the expected score).

[102] Note that here 'truth' refers only to the realization of the event E which is unknown to the external observer. We are not talking about or suggesting the existence of a true (but unknown) probability.

false.[103]

- The quadratic score discourages extreme probability assertions that go beyond the available evidence. Generally, having incomplete knowledge about the truth or falsity of an event of interest means that one's probability ought to be different from 1 and 0, respectively. If, nevertheless, one decides to make extreme probability assertions, the quadratic score will lead to maximum penalties if the event of interest does not turn out in the direction that one thought: for example, if one asserts probability $\alpha=1$, but the event of interest is false, then the maximum penalty of 1 (i.e., α^2) is obtained. If one would have been more cautious, say assert $\alpha=0.8$ instead of $\alpha=1$, then the penalty would be only $0.8^2=0.64$ if the event turns out to be false. Analogously, the penalty would also be 1 if one asserts $\alpha=0$ when in fact the event of interest is true. Here, too, a more cautious assessment, such as $\alpha=0.2$, would lead to a smaller score, i.e. $(1-0.2)^2=0.64$.

- A recurrent critique of subjective probability is that it tolerates 'random' probability assertions, so-called guesses, regardless of consideration given to where the actual truth lies. Under the quadratic scoring rule, however, such a decision behaviour (i.e., random guessing) is not recommended, for the following reason. Suppose that one randomly states either probability $\alpha=1$ or $\alpha=0$, leading to the maximal penalty of 1 if one's assertion does not correspond to the truth-value of event *E*. Then, in a series of *n* distinct probability assertions (i.e., for distinct events), if one gets it 'right' (i.e., assert $\alpha=1$ if *E* is true, and $\alpha=0$ if *E* is false) half of the time, the score incurred will be *n*/2. Now, compare this result with the score incurred by a person who is more prudent and prefers not to take sides in one way or another, and hence assert probability $\alpha=0.5$ in each case. This means that for each probability assertion, whether the event of interest turns out to be true or false, the score incurred will be $0.5^2=0.25$. In a series of *n* probability assertions, the overall score will thus be *n*/4, which is smaller than that of the 'random guessing' strategy. Hence, subjective probability assertions are not 'anything goes' exercises, but can be improved with careful reflection.

We have not seen, so far, how the above ideas favour sincere probability assertions. As an illustration, consider the example of tossing an unbiased coin. To be clear, coin tossing is clearly not of interest in legal proceedings. It is used here solely to ensure that there is no ambiguity about the uncertain event to be assessed (i.e., whether the outcome of the single toss is head or tails), so that attention can be directed to the nature of probability and its

---

[103] Indeed, for a probability assertion of $\alpha=1$, the penalty would be 0 if the event E is in fact true (analogously for a probability assertion of $\alpha=0$ for an event E that is false).

conceptualisation through scoring rules. More realistic illustrations will be mentioned later.

We discuss two aspects in turn: first, the meaning of an asserted probability resulting from the measurement against a standard (following arguments presented in §2.B. and second, consideration of the asserted probability under the quadratic scoring rule.

Regarding the first aspect, it is widely accepted that the probability for the outcome heads (or tails) after tossing an unbiased coin is 0.5,[104] because the way in which the tossing experiment is conducted gives us no reason to consider that one outcome will be favoured over the other. It is worth noting that this information – knowledge about the tossing mechanism at work and, possibly, knowledge about the outcomes of past tosses – is evidence, actually. It is one source of information for justifying a probability assertion of 0.5. Specifically, asserting that our uncertainty about the outcome of a toss turning out heads (or tails) is measured by a probability of 0.5 means that we hold the same uncertainty for the event of drawing a black ball from an urn that contains the proportion $p=0.5$ of black balls.

Having seen that our degree of belief about the outcome of the single coin toss is 0.5, we now turn to the second aspect, i.e. how to provide an incentive so that the asserted probability in fact corresponds to the actual probability of 0.5 we have in our mind. It is here that the scoring rule scheme intervenes. The question thus is, looking at the full range of possible probability assertions (i.e., the values between 0 and 1), how can we motivate the person to state 0.5, rather than a value other than the one she bears in mind (i.e., 0.5)? The point is that under a proper scoring rule, the *decision* to state 0.5 is optimal in the sense that it is the decision which minimises the *expected score*. The expected score is computed in two steps. First, if the coin falls on heads (i.e., the event E is true), our statement of probability 0.5 will lead to the loss $(1-0.5)^2=0.25$. The probability of incurring this loss is 0.5, i.e. our probability for the event E.[105] Therefore, we have $0.5x0.25=0.125$ as the first part of the computation of the expected score. Second, if the coin falls on tails (i.e., event E is false), our statement of 0.5 will also lead to a loss of $0.5^2=0.25$. Again, this loss will be incurred with probability 0.5, the probability of event E being false. Thus, we have a second term $0.5x0.25=0.125$ which contributes to the expected score. These two components combine by addition to give the expected score of stating 0.5, written ES(0.5):

---

[104] Note that we are only contemplating the outcome of a single toss here. We are not considering the relative frequency of heads or tails in a long series of repeated tosses.

[105] The important point at this juncture is that it is not necessary to know the ground truth in order for the scoring rule to be useful. The ground truth is only contemplated to define the score. It is to be multiplied by one's probability for the event of interest (here *E*).

$$ES(0.5)=0.5x(1-0.5)^2+0.5x0.5^2=0.25. \qquad (1)$$

The general formula for the expected score (ES) for an asserted probability α is

$$ES(\alpha)=Pr(E)x(1-\alpha)^2+Pr(not\text{-}E)x(\alpha)^2, \qquad (2)$$

where $Pr(E)$ is our personal (genuine) probability (i.e., degree of belief) for the event *E*, as defined above, depending on our knowledge, understanding and any relevant evidence available at the time our assessment is made. Note that the use of one's probabilities for the events *E* and *not-E* shows that actual ground truths are not necessary.[106] The core of the argument is the use of the probabilities for the events *E* and *not-E* to compute the expected score which serves as a measure to compare different belief statements α. In particular, note that the optimal probability α to be reported, i.e. the value α that minimises the expected score (Eq. (2)), is the one that corresponds to $Pr(E)$.

Let us provide some further illustration of this fundamental property of the quadratic scoring rule by showing that asserted values α different from $Pr(E)=0.5$ lead to higher expected scores, and hence are not optimal decisions. Consider the assertions of, for example, $\alpha=0.3$ and $\alpha=0.8$. Using Equation (2), the expected scores for these two probability assertions are, respectively:

$$ES(0.3)=0.5x(1-0.3)^2+0.5x0.3^2=0.29, \text{ and}$$
$$ES(0.8)=0.5x(1-0.8)^2+0.5x0.8^2=0.34.$$

To provide a complete picture, we can plot the expected score for all possible probability assertions α, as shown in Figure 1.
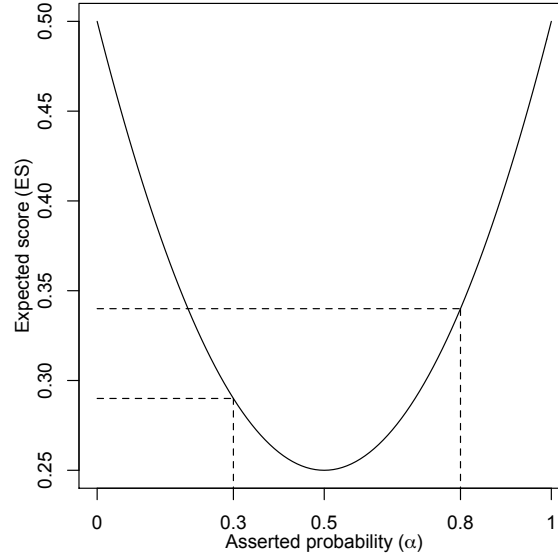
---

[106] See also supra note 105.

Figure 1: Representation of the expected score (ES) as a function of the asserted probability α when the probability for the event of interest, *E*, is Pr(*E*)=0.5. The optimal decision is to report the probability α that corresponds to one's probability Pr(*E*). All decisions to assert probabilities α different from 0.5 have higher expected scores. Examples (using dashed lines) are given for decisions to report α=0.3 and α=0.8, leading to expected scores of 0.29 and 0.34, respectively.

Note further that the principle that it is optimal to assert the probability α that corresponds to one's probability Pr(*E*) is valid for any probability, not just for the case of Pr(*E*)=0.5 considered in the example above. We can illustrate this graphically as shown in Figure 2. The x-axis in this figure represents the range of different degrees of belief, Pr(*E*), that one may have, in the range from 0 to 1. For each of these degrees of belief, there is exactly one asserted probability α that is optimal, i.e. α=Pr(*E*). Consider, for example, the decision to report probability α=1. The straight line from the left top corner to the bottom right corner shows the expected score for this decision as a function of one's actual belief state, Pr(*E*). Clearly, if one's belief is Pr(*E*)=0, the decision to report α=1 has the highest expected score and, thus, will be the worst decision to make. All decisions to report values α<1 have a smaller expected score. The decision to report α=1 is the best decision only if Pr(*E*)=1. In the latter case, the expected score is 0, and there is no other decision (i.e., report of a value α<1) that has a smaller expected score. More generally, we can plot the expected scores for any decision α – in Figure 2 examples are given for vales α={0,0.2,0.4,0.5,0.6,0.8,1} – to show that a reported probability α is optimal (i.e., has the minimal expected score) exactly when α=Pr(*E*), thus showing that under the chosen scoring rule

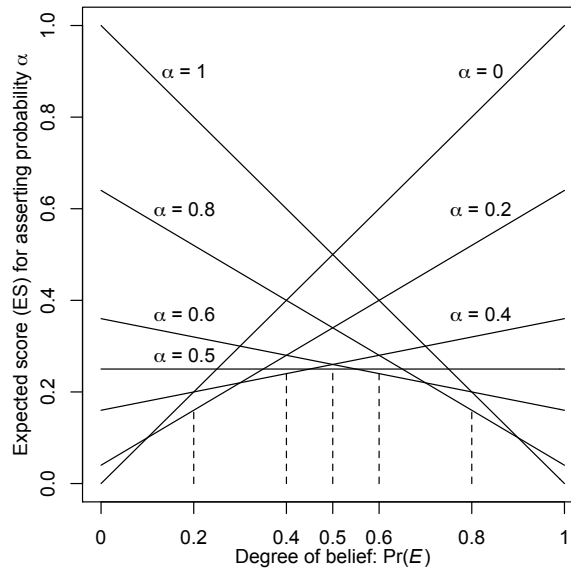reporting sincerely one's probability is always optimal.[107]

Figure 2: Representation of the expected score (ES) for different decisions (i.e., reported probabilities α), as a function of one's actual degree of belief, Pr(*E*), calculated according to equation (2). Examples are given for values α={0,0.2,0.4,0.5,0.6,0.8,1}, showing that reporting a value α is optimal – i.e. minimises the expected score – when the reported value α corresponds to one's actual degree of belief Pr(*E*).

C. Discussion of the decisionalist perspective to probability

At the outset, let us insist on the point that in the legal context, the scoring rule scheme is not a way to audit a probability assessor (e.g., an expert who reports a probability) in actual cases. Instead, it is an analytical argument that serves as a way to understand why genuine probability reporting is optimal. More generally, the scoring scheme illustrates the feasibility of conceiving incentives[108] to encourage people to assert personal probabilities honestly.

---

[107] For a representation with values α in 0.1 steps see, for example, de Finetti (2017), supra note 7, at 160.

[108] The idea of the scheme serving as an incentive may be challenged on the ground that people may not adhere, consciously or not, to the ideal of making up their mind carefully and reporting genuinely evidence-based scientific opinions. To the extent that this may occur in current legal practice, our analyses in this paper will rest on an idealistic starting point. As de Finetti has noted, "[it] is of course necessary to distinguish, above all, those people who are aware of the mechanism [i.e., the scoring rule] and the spirit of the device and who try to use it correctly to express their own beliefs from those who prefer, owing to a lack of right understanding or of willingness, to behave as gamblers, for example, by trying to attain the maximum of the scores by giving 100 per cent. to the most probable case. The latter must be studied separately, as a pathological class." (Bruno de Finetti, Does it Makes Sense to Speak of 'Good Probability Appraisers'?, in: Irving J. Good, Alan J. Mayne & John M. Smith

Under the presented scheme, its neither advisable to distort numbers, in one direction or another, from one's actual degree of belief, nor is it advisable to declare probabilities in a random way (i.e., by careless guessing). These insights should help tame critiques[109] that label subjective probabilities as aleatory concepts that are not tailored in any way to available evidence. Indeed, without a device such as a proper scoring rule, there is nothing in the rules of probability that – strictly speaking – would forbid us to assert any probability, in the sense of a 'random guess'. Insofar, concerns about the 'seriousness' of personal probability assertions are legitimate and deserve careful consideration.

But despite the thorough properties of scoring rules, critics may argue that the idea of scoring is of little practical value because one does not 'know', in the first place, what one's probability is (i.e., what probability one has in mind). In particular, so the objection, there is nothing to be contemplated about in terms of an 'optimal' probability α, to serve as a reported value, since there is no underlying degree of belief $\Pr(E)$ available against which α could be evaluated. This is a subtle observation for which different replies can be offered. First, it is important to note that the difficulty raised does not relate to the discussed scoring scheme, but refers to the difficulty of probability assignment in the first place. As we have noted repeatedly throughout this paper, this is a general difficulty that is not specific to subjective probability, but to any interpretation of probability – for as if there is any data and knowledge available, the subjectivist perspective will, as any other interpretation of probability, strive to use these data sources as suitably as possible. Second, the issue is not one of 'knowing' or not (one's) probability; in the subjectivist's perspective one knows, as per definition, one's probability, for it would not make sense to say that one does not know one's state of mind.[110] The question is if one is prepared to undertake the burden of pinning one's probability down,[111] at least in terms of an order of magnitude, and provide a justification for it; else the matter might better not be debated at all, at the price of offering no assessment at all.

But still, critics may say that given the number of occasions and the

---

(Eds.), The Scientist Speculates: An Anthology of Partly-Baked Ideas (1962), 357–364, at 352).

[109] See, for example, supra notes 48, 49 and 50.

[110] See also discussion presented in §I.E.

[111] Skeptical readers may say that this is not a problem, for example, in the field of forensic science where examiners tend to be happy to state their probability, especially extreme values of 1 and 0 *despite* the fact that the issue of interest is not one of certainty and impossibility, but marked by uncertainty. However, the deeper problem here may be a lack of scientific integrity (see also supra note 108) or a misunderstanding of what it means to have incomplete information that, by definition, is insufficient to warrant a categorical conclusion. Further discussion of this point is presented later in this section.

variety of cases where it *is* difficult to make up one's mind, the practical relevance of the scoring rule scheme is limited. This objection is formulated from a strongly applied perspective, and is similar to objections levelled against the measurement of probability using a standard. Clearly, as much as we do not suggest to carry out the measurement of uncertainty against a real, physical urn, we do not suggest that probability assessors compute expected scores. What is important, however, is the use of these ideas as conceptual devices. As regards the scoring scheme, it serves the purpose of clarifying that, ultimately, the task of asserting a probability comes down to a decision: the person *must decide* which value is the most defensible figure, given the best knowledge available to that person at the time when the probability assertion is made. This is not to say that the person shall not invoke symmetry considerations[112] or summaries of relevant data (such as relative frequencies[113]) to inform her probability assignment.[114] The point of the scoring scheme is to avoid that people defer to abstract and seemingly impersonal definitions of probability and, by doing so, think that they hold no responsibility for their probability assertions, especially in cases where their assessments would be found problematic or otherwise dubious.

Historically, the quadratic scoring rule is known in meteorology as the Brier score[115] and has been widely studied since around the 1970s, but it is also very relevant for experts consulted in legal proceedings. Forensic science experts, for example, are regularly called on to inform a fact-finder about the probability of observing particular features in recovered material or marks (e.g., the DNA profile of a crime stain, or a configuration of ridge skin characteristics in a fingermark) under the assumption that the material or mark comes from an unknown person, different from the person of interest charged by the prosecution. For single donor DNA profiles that are (nearly) complete, and good quality fingermarks, such probabilities can be extremely low, to the point that some experts are tempted to report that it is impossible

---

[112] Broadly speaking, symmetry judgments refer to situations in which one has no reason to prefer one outcome to another. Such judgments may be based on past experience (e.g., multiple observations of a given phenomenon) or one's knowledge about a particular physical process at work.

[113] Note, again, that we do not consider here relative frequency as an admissible definition of probability, but only as a summary of data suitable to *inform* a probability assignment (see also §I.A.).

[114] This includes considerations from probability definitions that, throughout this paper, we have identified as being deficient in various respects. Most interestingly, even de Finetti (2017), supra note 25, as a declared subjectivist, was not entirely dismissive of probability definitions that he did not consider sound: "(...) definitions of 'objective' probability, although useless *per se*, turn out to be useful and good as valid auxiliary devices when included as such in the subjectivistic theory." (at xvi)

[115] Glenn W. Brier, Verification of Forecasts Expressed in Terms of Probability, 75 Monthly Weather Review 1–3 (1950).

(i.e., probability of zero) to observe the evidential features in a person other than the person of interest – thus suggesting (overwhelmingly) strong support for the prosecution's proposition. However, asserting that the probability of seeing the evidential features in the event that an unknown person is the source is zero is both empirically unfounded and rationally unsound; it is a statement that goes beyond the available evidence.[116] Empirically, the statement is unfounded because the experts do not usually demonstrate that all other potential sources are in fact excluded, except in some atypical closed-set situations. And even if they presented evidence of comparisons with persons other than the person of interest, the factual accuracy of each individual comparison would hinge on probabilities of error not equal to zero (as human comparative examinations are known not to be error-free).[117] Hence, an expert's personal probability cannot be zero, and rounding it off to zero would, under a proper scoring rule, *not* be an optimal value to report. The scoring rule scheme thus is one way of showing experts why they should not distort probabilities away from their actual degree of belief.[118] Note that the probability assignment task here is a general one, encountered also with other types of evidence.

A further area of expert evidence where scientists should avoid the reporting of categorical opinions (i.e., certainty statements) involves DNA mixtures, i.e. traces that contain genetic material from more than one person. Based on the number of characteristics seen in the DNA mixture profile, some scientists seek to infer the number of persons who have contributed material to the mixed stain. Invoking ad-hoc criteria, such as the maximum allele count method, they make assertions about the minimum number of persons necessary to produce the set of characteristics seen in the trace. However, as real-life cases have shown, categorical statements about the number of contributors can lead to paradoxical results,[119] such as a particular person of interest being considered excluded or non-excluded as a potential contributor to a mixed stain, depending on the assumed number of contributors. One of the sources of such complications is the categorical assertion of a number (of contributors) that, in reality, is not and cannot be

---

[116] In a seminal paper, Stoney has called this a "leap of faith" (David A. Stoney, What Made us Ever Think we Could Individualize Using Statistics? 31 Journal of the Forensic Science Society 197–199 (1991), at 198).

[117] Alex Biedermann & Joëlle Vuille, Understanding the Logic of Forensic Identification Decisions (without numbers), sui-generis 397–413 (2018), https://doi.org/10.21257/sg.83.

[118] Alex Biedermann, Paolo Garbolino & Franco Taroni, The Subjectivist Interpretation of Probability and the Problem of Individualization in Forensic Science, 53 Science & Justice 192–200 (2013).

[119] See, e.g., Alex Biedermann, Franco Taroni & William C. Thompson, Using Graphical Probability Analysis (Bayes Nets) to Evaluate a Conditional DNA Inclusion, 10 Law, Probability & Risk 89–121 (2011).

known with certainty. A scientifically more honest report should acknowledge uncertainty and express it in terms of probability. In this context, a scoring rule helps to show that – conceptually – it is preferable to assign probabilities to different possible numbers of contributors, rather than assigning a probability of 1 to only one particular number of contributors (and zero to all other possible contributors). Further, probabilistic reporting not only avoids paradoxical results of the kind mentioned above, it can also empirically be shown – in terms of incurred scores over multiple assessments – to be superior to certainty statements (i.e., a categorical reporting format).[120]

Of course, the crucial question that remains is whether scientists are able to provide justified probability assignments for the situation at hand supported, ideally, by relevant data.[121] We entirely agree with this call for challenging assigned probabilities, as outlined in §I.C. but this aspect is not covered by scoring rules[122] – these only seek to encourage the honest reporting of probabilities within the scope of what is argumentatively defensible. Again, either one is prepared to address the burden of justification, and assume responsibility for any assigned value, or else no opinion should be offered.

D. Objectivism versus objectivity in subjectivism: "Nothing is lost what was a mere illusion"[123]

Our argument throughout this paper has been that there is no way for any person asserting a probability to avoid taking a personal stance. All attempts to conceive of some sort of 'objective' probability, independent of the person uttering a probability, are bound to fail because, by necessity, adopting a criterion or definition requires itself, at some point, a personal choice[124] – or

---

[120] For a demonstration using simulated DNA mixtures with known numbers of contributors see, e.g., Alex Biedermann et al., Inference About the Number of Contributors to a DNA Mixture: Comparative Analyses of a Bayesian Network Approach and the Maximum Allele Count Method, 6 Forensic Science International: Genetics 689–696 (2012).

[121] See Allen (2017), supra note 6, for a critical discussion of this point.

[122] More generally, challenging and scrutinizing reported probabilities in practice is a question of auditing and disclosure, which in turn is a problem of regulation.

[123] Bruno de Finetti, Bayesianism: its Unifying Role for Both the Foundations and Applications of Statistics, 42 International Statistical Review (1974) 117–130, at 121.

[124] Among philosophers of science, this view is also known as 'De Finetti's mouse-trap': *"Erzähle mir, welche Variante des Objektivismus du vertrittst, und ich will dir sagen, auf welchem Weg du in de Finettis Mausefalle gelangst!"* (italics as in original) (Wolfgang Stegmüller, Das Problem der Induktion: Humes Herausforderung und moderne Antworten (1975), at 27). This can be translated as: "Tell me what variant of objectivism you support, and I will tell you in what way you fall into De Finetti's mouse-trap." Here Stegmüller argues that one cannot avoid to admit, at some point, that one is presupposing the notion of personal probability. Similarly, Howson & Urbach, supra note 16, argue that "(...)

an interpersonal agreement.[125] For virulent critics of subjective probability, in both science and the law, this position might be offensive and inacceptable, for they regard objectivity as the ultimate virtue to which all human activity ought to aspire.[126] Underlying these concerns is a rather stereotypical distinction between what is objective and what is not, and hence is, necessarily, unacceptably subjective – in the sense that a subjective statement could not, by definition, have any objective status.[127] But the matter is more subtle than that. A brief comment thus is in order.

The opposition between subjectivism and objectivism is not a recent observation.[128] It was also at the centre of attention for the founders of subjective probability, such as de Finetti. Interestingly, the subjectivist de Finetti was not dismissive of objective notions, but insisted on carefully defining their scope. For example, he admitted that "(...) statements have objective meaning if one can say, on the basis of a well-determined observation (which is at least conceptually possible), whether they are either TRUE or FALSE".[129] More generally, objectivity in probability *assessment*

---

the ideal of total objectivity is unattainable and that the classical methods, which pose as guardians of that ideal, in fact violated it at every turn; virtually none of those methods can be applied without a generous helping of personal judgement and arbitrary assumption" (at 9).

[125] Even statisticians, whose mathematical methods might seem most suitable for 'objective' approaches, clearly reject objectivism. Dennis V. Lindley, That Wretched Prior, 1 Significance (2004), has noted: "Objectivity is merely subjectivity when nearly everyone agrees" (at 87); and for José M. Bernardo & Adrian F. M. Smith, Bayesian Theory (2000) "(...) objectivity has no meaning in this context apart from that pragmatically endowed by thinking of it as a shorthand for subjective consensus" (at 424).

[126] The quote in the title of this section is part of the following excerpt from de Finetti (1974), supra note 123: "Such ideas are however distressing for some people, who consider objectivity, in the strictest sense, as a necessary attribute of probability and of science. But regret for losing the faith in the perfect objectivity of probability, and hence of science, is unjustified. Nothing is lost but what was a mere illusion" (at 121).

[127] As argued by de Finetti, the opinion that well-defined notions must have an objective meaning is ill-conceived for as subjective notions can also be well-defined (Bruno de Finetti, Le Vrai et le Probable, 3 dialectica 78–92 (1949), at 81).

[128] Popper, for example, has noted: "The words 'objective' and 'subjective' are philosophical terms heavily burdened with a heritage of contradictory usages and of inconclusive and interminable discussions." (Karl Popper, The Logic of Scientific Discovery, 2005, at 22)

[129] Supra note 7, at 5. De Finetti expressed the same idea previously as follows: "The only probability assessment that is 'objectively fair' or 'objectively privileged' or that 'belongs to the field of objective reality' (...) would be that which would attribute *probability 1 to the events that will take place and the probability 0 to others*" (De Finetti (1949), supra note 127, at 91) [italics as in original, translation by the authors; original: "La seule évaluation de probabilité 'objectivement juste' ou 'objectivement jouissant d'une situation de privilège' ou 'appartenant au domaine de la réalité objective' (...), serait celle qui attribuerait *la probabilité 1 aux événements qui auront lieu et la probabilité 0 aux autres*."]

clearly was a concern for de Finetti, as he was mindful of how to meaningfully and justifiably assign probabilities, in the light of available evidence and knowledge. His decisional approach to probability assertion, using the method of scoring rules was his main instrument to approach this challenge.[130] It is important to note that in this scoring rule scheme, the notion of 'goodness' of probability assertions does *not* relate to a supposed objective probability. Instead, 'goodness' is assessed by comparison to the truth-values of the uncertain events and, as mentioned above, de Finetti did not question that those truth values *are* objective.[131] However, he opposed the idea of *objectivism*[132] by which he referred to claims that ascribe to probability a meaning that is independent of the thinking individual – i.e. the existence of probability as a property of the outside world.

In summary, thus, the genuine personal probabilities – in the de Finettian subjective sense – *do* and must, as suitably as possible, acknowledge and incorporate objective elements. This brings us back to de Finetti's insistence on distinguishing between the *definition* of probability and the assignment of a probability in an individual case. Regarding this distinction, Carla Maria Galavotti, in her foreword to De Finetti's philosophical lectures on probability, explains:

> "Unlike subjectivists, that do not confuse the definition and the evaluation of probability, upholders of the other interpretations do; they embrace a unique criterion – be it frequency or symmetry – and use it as grounds for both the definition and the evaluation of probability. They take a 'rigid' attitude towards probability, which implies a univocal choice of the probability function. On the contrary, the subjectivist approach is 'elastic', it does not commit the choice of one particular function to a single rule or method, but regards all coherent functions admissible."[133]

This versatile aspect of probability should look attractive for legal

---

[130] In his introductory lesson of a series of lectures held at the Institute of Advanced Mathematics in Rome (Italy) in 1979, de Finetti said: "I have asserted that the only serious device to talk about probability is the one based on proper scoring rules" (De Finetti (2008), supra note 25, at 9). Earlier traces of this view can be found in the 1960s, e.g., de Finetti (1962), supra note 108.

[131] Alberto Mura (2008), supra note 25, Editor's Notes, at 13.

[132] For a detailed discussion de Finetti's distinction between objectivity and objectivism, and references to his original works in Italian, see Galavotti (2018), supra note 98, at 153, and Maria Carla Galavotti, Subjectivism, Objectivism and Objectivity in De Finetti's Bayesianism, in: David Corfield & Jon Williamson (Eds.), Foundations of Bayesianism, Applied Logic Series (2001), 161–174.

[133] Carla Maria Galavotti, De Finetti's Philosophy of Probability, in: Bruno de Finetti (2008), supra note 25, at xx.

scholars who emphasize a pluralistic perspective to belief formation.[134] They should also find value in the fact that, as Galavotti further explains, (de Finettian) subjectivism is an appeal to the careful appraisal of the multitude of sources of relevant information that one encounters, a view quite contrary to the misinterpretation of subjective probability as 'anything goes' opinions:

> "For de Finetti, the evaluation of probability is the result of a complex procedure involving both subjective and objective elements. Indeed, for him the explicit recognition of the role played by subjective elements within the formation of probability judgments is a prerequisite for the appraisal of objective elements. He always warned that the objective component of probability judgments, namely factual evidence, is in many respects context-dependent: evidence must be collected carefully and skilfully [sic], its exploitation depending on the judgment on what elements are relevant to the problem under consideration and enter into the evaluation of probabilities. Furthermore, the collection and exploitation of evidence depend on economic considerations that vary according to the context. For all of these reasons, the exploitation of factual evidence involves subjective elements of various kinds."[135]

Ultimately, for de Finetti, this comprehensive perspective provides a decisive advantage for the subjectivist over the objectivist views, not least because – by their attempt to define single and unique criteria – objectivist approaches must necessarily ignore some aspects to the detriment of the value that these aspects could add, thus depriving evaluators from relevant information.[136] The genuinely subjective account, instead, is always free to exploit the entirety of the thinking mind's cognitive means and sources of information.[137] In one of his many forceful commentaries, de Finetti has

---

[134] See, e.g., Allen (2015), supra note 85: "Fact finders come to trial with a vast storehouse of knowledge, beliefs, and modes of reasoning that are necessary to permit communication to 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." (at 79); and similarly, Ronald J. Allen, *The Declining Utility of Analyzing Burdens of Persuasion*, 48 Seton Hall Law Review 995–1015 (2018): "fact finders employ their full panoply of cognitive tools" (at 998).

[135] Carla Maria Galavotti, supra note 133, at xx–xxi.

[136] The reductionist tendency of objectivist views is precisely the feeble point that critics of probability in the law have largely, uncompromisingly and, in our view, rightly exploited and exposed.

[137] In an overview of de Finetti's works, Ludovico Piccinato, *De Finett's Logic of Uncertainty and its Impact on Statistical Thinking and Practice*, in: Prem K. Goel & Arnold Zellner, Bayesian Inference and Decision Techniques, Essays in Honor of Bruno de Finetti, (1986) 13–30, has noted: "The inferential theory of de Finetti relies on some cornerstones which have a clear conceptual meaning, but the theory itself does not imply a complete and

expressed this view as follows:

> "Not being able to invent methods of synthesis comparable in power and insight to those of the human intuition, nor to construct miraculous robots capable of such, he [the objectivist] contents himself, willingly, with simplistic schematizations of very simple cases based on neglecting all knowledge except a unique element which lends itself to utilization in the crudest way."[138]

Again, this appeal to a person's entirety of knowledge is compatible with how legal scholars view reasoning processes at trial, that is organic and operated by humans.[139]

As a last and contemporary illustration of the inevitability of personal probabilities in the de Finettian sense, consider probabilities used in the context of forensic DNA evidence. In this field, figures are often referred to as objective (or, 'statistical') probabilities by both scientists and non-scientists alike because – unlike in other forensic fields where relevant datasets are sparse – abundant genetic data are now widely available in specialist literature. However, as emphasized by Evett et al., there is nothing inherently objective in DNA profile probabilities assigned on the basis of such data:

> "There appears to be a fairly widespread misconception that there is a real "statistical probability" to be assigned to a [DNA] profile but this is not the case. There is an infinite range of ways of carrying out the calculation that underlies the figure given. The method chosen in the individual case must be seen to be as much a matter of opinion as one given in other areas of forensic science. The match probability is "personal". It is based on what the scientist considers to be the most appropriate calculation given the circumstances of the case."[140]

E. Corollaries for probability assertions and their scrutiny

Though not unimportant, the debates over subjectivism and objectivism

---

detailed system of procedures, similar to the current *cookbooks* of the standard literature." (at 13)

[138] De Finetti (2008), supra note 25, at 7.

[139] E.g., Ronald J. Allen, On the Significance of Batting Averages and Strikeout Totals: A Clarification of the "Naked Statistical Evidence" Debate, the Meaning of "Evidence," and the Requirement of Proof Beyond a Reasonable Doubt, 65 Tulane Law Review 1093–1110 (1991) ("information content will be determined by a live human being exercising judgment organically rather than mechanically" at 1102). See also supra note 134.

[140] Ian Evett, Lindsey A. Foreman, Graham Jackson & James A. Lambert, DNA Profiling: a Discussion of Issues Relating to the Reporting of Very Small Match Probabilities, Criminal Law Review 341–355 (2000), at 346.

represent a marginal topic with little prospect of new arguments, mainly because the two notions are intertwined (with the latter being part of the former). More important are aspects such as honesty on the part of the person conducting an assessment, completeness of evidence and the elicitation of statements reflecting faithful representations. These aspects are particularly relevant in the context of expert witnesses testifying as to probability assignments. The decisionalist perspective to personal probability exposed throughout this paper lends itself to several remarks thereto.

Truthfulness, for example, is widely considered as the ultimate virtue, but empirical means to pursue it are rather limited. Oath, for example, is at best an encouragement or an appeal to honesty, but no guarantee.[141] The scoring rule scheme, too, might seem inapplicable at first sight because the recipient of a probabilistic assertion, prima facie, has no direct means to 'know' whether the asserted probability coincides with the person's actual degree of belief. All that the external observer can hope for is that the person *will* act in her interest by honestly stating her probability – reporting this value is the optimal decision, as we have seen, because it is the one with the minimum expected score. What is more, since the ground truth is not known in a criminal trial – for otherwise there would be no litigated matter and no need to consult an expert or witness – there is no way to know how far from the truth a given probability assertion is.[142] This does not mean, however, that a dishonest person has nothing to fear. Because her assertions expose her directly to her peers, she should care about the reasonableness of her opinions, and their faithful statement. Conceptually, such sincere reporting of scientific opinions is precisely the aim of considering probability assertions as decisions under a scoring rule scheme. Thus, an expert who intends to proceed responsibly and who wishes to avoid appearing unreasonable, should abstain from reporting an aleatory opinion.[143] She should think carefully about what a competent expert's opinion ought to be in the case at hand. This is an idealistic view of expert testimony that could face challenges such as partisan bias,[144] and that is not immune to possible expert disagreements.[145] Thus, in practice, procedural elements such as cross-

---

[141] E.g., Samuel R. Gross, Expert Evidence, 1991 Wisconsin Law Review 1113–1232 (1991), at 1178.

[142] Indeed, as de Finetti (2008) has noted supra note 25: "If we knew where truth was, we would not get close to it, we would go straight there. And if we do not know where it is, we cannot even know how far it is from us." (at 53)

[143] Recall that under the scoring rule scheme (§II.B.), reporting any value other than the one corresponding to one's actual degree of belief is suboptimal.

[144] For a critical analysis with references back to, among others, Wigmore, see e.g. George C. Harris, Testimony on Sale: The Law and Ethics of Snitches and Experts, 28 Pepperdine Law Review 1–74 (2000).

[145] As mentioned by Jackson et al. (2003), supra note 96, at 29, disagreement is not

examination and rebuttal require additional means for peer-reviewing and rationalizing the reporting of probabilities.

Following discussion presented earlier in §I.C. probability assertions ought to be accompanied by a justification, or an argument, because the recipient of expert information shall be able to understand how a particular opinion came to be, and on what evidence it is based. But there is a further dimension to this requirement. The recipient of expert information should also expect assistance in placing the expert's statement within a wider context, by which we mean an explanation by the expert of how her opinion compares to that of other qualified witnesses in her field of expertise. This is a subtle means to probe the comprehensiveness of the expert's knowledge and expertise in her field of specialty, or at least represents an incentive for her to self-critically assess the frontiers thereof. This is particularly crucial when the expert considers that her opinion is not in line with the view of peers in her field; it is important then to elicit the extent to which her opinion differs from mainstream views, and why (e.g., the expert may have more specialist knowledge and expertise than the average of experts in her field, maintain a different interpretation of available data or the state of the art, etc.).

Inquiring about the foundations of expert opinions is important because, as we have explained, the liberal concept of probability as such does not prescribe any particular probability values, nor is there any 'true' probability.[146] Thus, experts should consider it to be their duty to provide accompanying explanations, so as to help recipients of expert information absorb complex matters of scientific expertise, and judge their reasonableness.[147,148] Where such information is not provided by experts, opinions should be challenged so as to elicit the required information.

PART 3: DECISION STRUCTURES IN JURIDICAL EVIDENCE AND PROOF

A. Conceptual insufficiency of purely probabilistic standards of decision

In scholarly discussions of evidence and proof standards, references to probability abound, used either as a qualitative verbal notion to express a

---

necessarily a drawback because it can stimulate clarifying discussion. We provide further support for this argument in the paragraph that follows below.

[146] Supra note 47.

[147] Regarding the requirement of additional explanation, David H. Kaye, Clarifying the Burden of Persuasion: What Bayesian Decision Rules Do and Do Not Do, 3 International Journal of Evidence & Proof 1–28 (1999), has argued: "(...) no mathematical result is self-applying, and additional argument is necessary to bridge the gap from a general mathematical truth to a substantive application." (at 27)

[148] This is not to be understood as a disguised *argumentum ad populum*. The fact that an opinion is to some extent representative of a relevant group of people, such as scientists, is not – as such – a credential that merits authority, but is a feature that decision makers may wish to know about.

person's degree of uncertainty, or quantitatively as a minimum threshold required to reach a particular conclusion. Although there appears to be some consensus, at least in standard textbooks, on how expressions such as "preponderance of the evidence" (or "balance of probabilities") in civil cases and "beyond reasonable doubt" in criminal cases might be expressed in quantitative probabilistic terms, ample empirical research shows that there is diversity in how different individuals, from different backgrounds (e.g., prospective jurors, judges, citizens), actually associate numbers and words.[149,150] The difficulty appears to stem from two intertwined aspects: the first is the question of what is *the* exact probability threshold that ought to be reached in a given case, and the second aspect relates to the issue of 'knowing' when exactly one has reached a particular probability. Throughout this paper, we have focused on the latter aspect, arguing that while – conceptually – one's probability can be defined clearly, the nature of probabilities, their elicitation and assertion in practice can be further clarified and scrutinized through a decisional perspective. We shall now argue that this decisional perspective is also relevant for the former aspect, i.e. standards of decision. Specifically, we shall argue that traditional discourses about abstract probability thresholds are short-sighted in the sense that they misconceive the scope of the application of probability.

To illustrate our view, let us take a closer look at the attempts to define or assert a specific probability (threshold) for a particular *decision* (i.e., conclusion or verdict). By definition, as we have argued, an asserted probability only expresses the degree of belief that a given person entertains about the truth or otherwise of a particular uncertain event (or, proposition), based on all the knowledge, data and evidence available to that person at the time she asserts her probability. Of course, there is nothing in the theory of probability that either prevents or encourages the *definition* of an 'if–then rule' of decision of the kind 'if the probability of a particular proposition (e.g., Mr. A is the person who assaulted Mr. B) is greater than *x*, then decide so and so'. Legal literature abounds with arguments for the view that such rules may be unsuitable. Without going into the details of those objections, we agree

---

[149] E.g., Francis C. Dane, In Search of Reasonable Doubt: a Systematic Examination of Selected Quantification Approaches, 9 Law and Human Behavior (1985) 141–158; Jon O. Newman, Quantifying the Standard of Proof Beyond a Reasonable Doubt: a Comment on Three Comments, 5 Law, Probability & Risk 267–269 (2006); Rita J. Simon & Linda Mahan, Quantifying Burdens of Proof: a Review From the Bench, the Jury, and the Classroom, 5 Law and Society Review 319–330 (1971); Richard Seltzer, Russell F. Canan, Molly Cannon & Heidi Hansberry, Legal Standards by the Numbers, Quantifying Burdens of Proof or Search for Fool's Gold?, 100 Judicature 56–66 (2016).

[150] This is also a well-known observation in psychological literature and applied contexts such as military and intelligence analyses. For a discussion and further references see, e.g., Winterfeldt & Edwards (1986), supra note 96, Chap. 4.2.

that crude 'if (and only if)-then' rules are unsuitable to make decisions for the simple reason that, strictly speaking, probability – in isolation – is not designed to warrant particular decisions. It is important, however, to be precise about what to conclude from this observation, and it is here that our view diverges from the criticism commonly leveled at legal probabilism. The standard critique is that probability as a whole is unsuitable for decision-making in the law. In our view, however, it is a misconstruction to think that probability *alone* could logically warrant a particular decision; the criticism (of unsuitability) thus is essentially pointless because it blames the theory for something that it does not purport to do. The real problem, as we see it, is the failure of discussants to address the issue of standards of proof for what they are: questions of decision. Probability, as such, is only concerned with capturing and formally expressing uncertainty. It is a necessary but insufficient preliminary for making a decision. So rather than criticizing probability as such, attention should be drawn to the question of how to establish coherent uncertainty expressions – in terms of probability – as integral parts of formal procedures that lead to decisions. Supporters of the relative plausibility account of legal proof have worked on this aspect by discussing the role that probability may assume as one concept among others that is used in arriving at legal conclusions.[151]

To some extent, it is surprising that discourse about abstract probability thresholds, and attempts to use them as unique decision criteria,[152] are even entertained, because it is generally agreed that legal decisions involve many dimensions for which decision makers may seek optimisation.[153] For example, it is widely argued that upon making a decision (e.g., rendering a verdict), one ought to think about the possible consequences of the decision and their relative merit or inconvenience, notions more commonly referred to as stakes. Such aspects are relevant in the law[154] as much as in other contexts of societal, professional or personal decision-making. This is not to say that probabilistic decision thresholds cannot be defined; actually, they can be defined and there is a formal framework for this that allows one to take into account decision consequences and their relative (un-) desirability.[155] While

---

[151] Allen (2017), supra note 6. Note, however, that the relative plausibility account provides a distinctive solution to the question of standards of proof by considering related burdens to be ordinal rather than a cardinal (e.g., Allen (1991), supra note 139, at 1096).

[152] Note that a similar observation has been made about objectivistic attempts to devise a single criterion for probability assignments (§II.D.).

[153] Allen (2015), supra note 85.

[154] An often-cited example is Justice Harlan's concurrence to In re Winship 397 U.S. at 370–371.

[155] The framework for the coherent formal processing of probabilities and utilities (or, losses) is decision theory. Its first detailed articulation for application in the law is widely attributed to Kaplan (1968), supra note 23.

uncontroversial from a purely mathematical point of view, the application of this framework – especially the interpretation of the framework's elements with respect to particular aspects of legal cases – has caused considerable debate among practitioners and academics.[156] We will not reiterate this debate in full detail, nor will we take particular sides with one or the other of the pre-eminent positions.[157] Instead, we will focus on what we call *decision structures*, which are based on selected elements of the formal (decision-theoretic) framework that we consider capable of providing a conceptual and structural account of inference and decision problems in the law and (forensic) science. We will do so by adopting a historical perspective based on works of, among others, Bruno de Finetti. His views on the meaningful use of decision theory appear to be rarely acknowledged in the legal literature. We will point out how they can help construct arguments that do not raise the latent criticisms of decision theory in the legal context, and place the seemingly concurrent notions of probability and (relative) plausibility into perspective.

B. *The inevitable elements of decision structures and their interpretation*

1. *Elements of decision structures*

More than half a century ago, researchers from various fields engaged in the development of a lively area of investigation into questions of decision-making where the consequences of decisions typically have monetary consequences and the aim is to elaborate optimal decisions with respect to well-defined criteria.[158] In the decades that followed, mathematicians and statisticians, economists, psychologists and philosophers of science, among others, pursued decision-science research from prescriptive, descriptive and normative perspectives.[159] The early contribution by legal scholars to formal

---

[156] For a careful introduction see, e.g., David Kaye (1999), supra note 147, and for a critical discussion Richard S. Bell, Decision Theory and Due Process: A Critique of the Supreme Court's Lawmaking for Burdens of Proof, 78 The Journal of Criminal Law and Criminology 557–585 (1987) for a critical discussion of the practical applications of Kaplan's (1968), supra note 23, decision theoretic account.

[157] Some commentators even consider the debate to have exhausted itself (e.g., Allen (2018), supra note 134).

[158] E.g., Herman Chernoff & Lincoln E. Moses, Elementary Decision Theory (1959), Irving J. Good, Rational Decisions, 14 Journal of the Royal Statistical Society, Series B (Methodological) (1952) 107–114 (mentioning applications in the law at 112–113) and Robert Schlaifer, Analysis of Decisions Under Uncertainty (1969).

[159] E.g., A Science of Decision-Making, The Legacy of Ward Edwards, Jie W. Weiss & David J. Weiss (Eds.) (2009); Advances in Decision Analysis, From Foundations to Applications, Ward Edwards, Ralph F. Miles, Jr. & Detlof von Winterfeldt (Eds.) (2007); Johnathan Baron, Thinking and Deciding, Fourth Edition (2008).

decision-oriented research is illustrated by the work of Kaplan.[160] His work is remarkable and still a highly recommendable read today because of his capacity to express complex matters in accessible terms. Kaplan described the formal decision-analytic viewpoint as follows:

> "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 uncertain data that inconclusively support or discredit differing hypotheses about the state of the real but nonetheless unknowable world."[161]

As is clear from this statement, the problem is essentially one of decision making under uncertainty, i.e. choosing a course of action when one does not know which state of the world actually holds, though one may have some – most often incomplete – evidence about it. These elements are fundamental and inevitable in the sense that they are encountered by any decision maker and cannot be discussed away.[162]

Matters become more intricate and controversial in discussions about the application of the theory, especially the quantification of uncertainties using (subjective) probabilities, the valuation of consequences (through utilities or losses) and the combination of these elements through computations of expected value; i.e., expected utility (or, loss). Expected value provides a measure that allows one to compare competing decisions and provide the basis for decision criteria, such as the maximisation (minimsation) of expected utility (loss). The mathematical structure of these notions is not controversial. However, debates arose over what exactly these notions can add to the understanding and practice of proof processes. The extensive exchanges between Professors Ronald J. Allen and David H. Kaye provide a detailed account of the intricate matters and perspectives involved. Professor Kaye, for example, places an emphasis on the capacity of the theory to serve as a logically rigorous *analytical* method.[163,164] Though not denying this feature, Professor Allen focuses on the *descriptive* capacity of theories to provide satisfactory accounts of the actual working of the legal process and

---

[160] Kaplan (1968), supra note 23.

[161] Supra note 23, at 1065.

[162] Note that *not* making a decision is also a decision.

[163] E.g., Kaye (1999), supra note 147.

[164] Note that, for example, decision theory postulates coherence relationships between probabilistic beliefs about the truth of competing hypotheses and preferences among decision consequences. Specifically, it is possible to show that, given suitable assumptions, decision theory leads to the probabilistic decision threshold of 0.5 for the simple civil case, and higher thresholds for other cases (e.g., criminal). For discussions of particular thresholds see, e.g., Friedman (1997), supra note 24, Friedman (2017), supra note 16, and Edward K. Cheng, Reconceptualizing the Burden of Proof, 122 The Yale Law Journal 1254–1279 (2013).

the individuals involved, especially what they are actually able to do, in terms of their mental (unaided) computational capacities. Regarding this aspect, he considers decision theory and expected value computations as unsuitable.[165] These disagreements concern the implications of particular results when applying decision theory in its full force.[166] And yet, the problem is almost universally considered – even among critical commentators of mathematical approaches[167] – as one of decision under uncertainty. No one denies that decisions *must* be taken,[168] that decisions have consequences of varying desirability (i.e., accurate and erroneous decision outcomes), and that we usually are uncertain about which consequence exactly will follow our action. The fundamental point is that we can identify these elements at various stages of action in the legal process and it is for this reason that we refer to those elements as forming fundamental *decision structures*, or the *anatomy of decisions*. Our main argument in this paper is that this *decisional nature* of particular aspects of legal proof is of wider significance as it can also serve to critically analyse the notions of probability and relative plausibility. We shall now provide a sketch of these decision structures to illustrate the commonalities of their underlying tenets.

2. *General decision structures in verdicts and assertions of probability and relative plausibility*

We start out by considering the most widely recognised decision structure, concerning ultimate issues. A trier of fact is presented with one side's (e.g., plaintiff or prosecution, P for short) allegation, and an opponent's (i.e., defendant's, D for short) contradiction of that allegation. An example for a civil case could be a claim in which the plaintiff seeks to prove that he suffered an injury as a result of a negligent act by the defendant. Decisions in favour of one side or the other may lead to accurate or erroneous outcomes (i.e., decision consequences), depending on whether or not P's or D's allegation is true. Accurate decision consequences – i.e. factually correct findings – are preferred to erroneous findings, and possibly there is also a preference ordering among the erroneous findings (e.g., a false conviction in a criminal case is usually considered worse than a false acquittal). The anatomy of this decision problem thus consists of verdicts (i.e., decisions or possible courses of action), states of nature (i.e., the parties' allegations or

---

[165] E.g., Ronald J. Allen, The Error of Expected Loss Minimization, 2 Law Probability & Risk 1–7 (2003).

[166] Friedman (1997), supra note 24, has made the same observation regarding the application of probability and Bayes' theorem ("the theory need not be applied in its most powerful gear", at 288).

[167] E.g., Allen & Pardo (2018), supra note 5, ("virtually all (probably all, actually) juridical fact finding is done under conditions of uncertainty", at 9).

[168] See also §I.D. and supra note 67.

'stories') and decision consequences.

Keeping the logic of this decision structure in mind, we can now see that there is no difference, in principle, with the idea of *probability as a decision* presented in §II.B. [169] A person required to assert a probability α regarding an event *E* must *decide* on a particular value and do so in a state of incomplete knowledge about whether or not *E* is true. As shown in Figure 3, read from left to right, the consequence of asserting probability α will result in a distance with respect to the truth-value of event *E*. This distance is zero if one asserts α equal to 1 (0) when *E* is true (false). The desirability of the various outcomes is expressed through a positive score (shown on the far right-hand side in Figure 3). This score is given by the square of the difference between the asserted probability α and the actual truth-value of the event *E*. The decomposition of probability assertion as shown in Figure 3 is known as a decision tree. Note that there is a further dimension to decision tree analysis known as 'folding back'. It consists in computing the expected score (ES) for each decision α,[170] obtained by weighing the score of each decision outcome by its probability of occurrence, and summing the products thus obtained. As outlined in §II.B. the optimal decision α, i.e. the decision with the minimum ES, is the one that corresponds to one's probability of the event *E*. We shall not repeat the full development of this computational level at this point because it goes beyond the descriptive exposition of the fundamental decision structures.
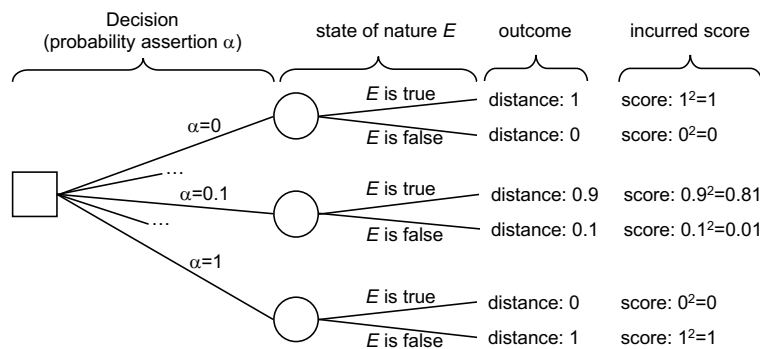

Figure 3: Decision structure for asserting a probability α for an event of interest, *E*, which

---

[169] See Michael J. Saks & Samantha L. Neufeld, Convergent Evolution in Law and Science: the Structure of Decision-Making Under Uncertainty, 10 Law, Probability & Risk 133–148 (2011) for a comparison between legal decision-making and scientific (i.e., classic) hypothesis testing. The authors find that there is little cross-fertilization between the two fields. This is unsurprising, given their restricted focus on frequentist hypothesis testing, leaving aside the decision-theoretic account that we consider in this paper.

[170] §II.B. Equation (2).

may either be true or false. The difference between the asserted probability α and the truth-value of *E*, i.e. the decision outcome, is valued in terms of a score given by the squared difference between the asserted probability and the truth-values 1 and 0, respectively, of event *E*. Note that, formally, this representation is a partial decision tree because probabilities for states of nature and computations of expected scores are not included.

The structural analogy between the above examples is insightful because, prima facie, decisions concerning ultimate issues and the assignment of a probability for an uncertain event may appear to be quite different tasks. Notwithstanding, it is clear that – analytically – conclusions in favor of the plaintiff's (P) and defendant's (D) accounts can be represented in a tree-like form of the kind shown in Figure 3, reduced to two decision branches. In turn, the real-world states would be the actual truth of either P's or D's account, leading to accurate and erroneous findings for P and D, respectively, depicted on the far right-hand side of the decision tree.

These analogies raise the question of how the decisional view of key operational tasks in legal processes relates to other perspectives, such Allen and Pardo's relative plausibility account of the structure of juridical proof. At its core, their account amounts to evaluating the parties' allegations through considerations such as coherence, completeness and consistency, and engaging in a deliberative process of assessing the congruence of the presented evidence with the competing versions of the contested events, using the fact-finder's entirety of knowledge, background information and cognitive capacity.[171] In terms of its structural elements, this account appears to involve the same elements as the general decision structure for ultimate issues sketched above:[172] a decision needs to be made for one of the competing stories, yet this decision is made in light of incomplete knowledge about the actual state of the world, thus leading to uncertainty about the correctness of the decision being made.

In broad terms, the relationship between relative plausibility and the decisional perspective may be perceived in at least two ways. One possibility is to consider relative plausibility as the argumentative background, justification and support for a fact-finder's selection of one of the available decisions (i.e., the finding for P or D, respectively). According to this view, the fundamental decision structure regarding ultimate issues is the one presented at the beginning of this section, with relative plausibility providing the explanation of the fact-finder's decision. A second possibility is to consider relative plausibility judgments as *decisions* in their own right. This view appears quite compatible with how Allen and Pardo circumscribe

---

[171] E.g., Allen (2017), supra note 6, at 140.

[172] To some extent, this is little surprising, since relative plausibility is specifically designed to capture, descriptively, decision-making regarding ultimate issues.

relative plausibility theory. Specifically, they note:

> "The "relative plausibility" theory explains the process as involving reasoning put to the effort of *deciding* the relative plausibility of the parties' various explanations of the evidence and contested events." [emphasis added][173]

Thus, understanding the identification, selection and assertion of the most plausible account as a decision under uncertainty can be graphically conveyed in terms of the decision tree shown in Figure 4. Note again that we focus only on a qualitative representation of the three essential elements that the decision maker faces: an exhaustive list of decisions (among which one element must be chosen), states of nature (though it is unknown which state of nature is actually true)[174] and decision consequences (outcomes). The decision tree represents these elements in natural flow, from left to right, thus providing an anatomy of the decision problem.
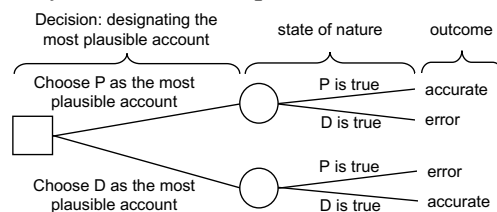

Figure 4: Representation of the designation of the most plausible account as a decision under uncertainty.

So far, our analysis has essentially been descriptive.[175] Besides, we have left the quantitative part of decision trees intentionally aside because – historically – it involves aspects over which legal scholars disagree. Instead, our aim was to delineate the inevitable structural aspects with which, hopefully, most discussants would agree. In this sense, the decision tree representation helps clarify where exactly current debates align and diverge. We work out these distinctions in some further detail below.

3. Reconsidering the interpretation of decision-theoretic advice

Proponents of the full decision-theoretic perspective would find the decision tree shown in Figure 4 incomplete because it does not provide an explicit criterion for selecting one particular course of action over an

---

[173] Allen & Pardo (2018), supra note 5, at 6.

[174] As an aside, note that if there were no uncertainty about the state of nature, there would be no decision problem in the first place: one would directly select the decision that leads to an accurate outcome.

[175] Also, we have restricted our analysis to a simple civil case. The relative plausibility reasoning scheme for the criminal case is more elaborate and goes beyond the scope of this paper. It is part of the authors' ongoing research.

alternative. Developing the decision tree in its complete form would require the assignment of utilities (or losses) for decision outcomes (at the far right-hand side of the tree) and of probabilities for the possible states of nature. Proponents of decision theory would use these elements to compute the expected utility (or loss) for each decision (i.e., assertion of the most plausible account), and then decide based on consideration of the decision with the maximum (minimum) expected utility (loss).[176] However, the framing of legal decision criteria based on the notion of expected value is a contested topic.[177] It is worthwhile, thus, to devote some discussion to it. It is at this point that we introduce a critical reflection on decision-theoretic guidance from a historical perspective.

Classic decision theory identifies the best decision as the one which maximizes (minimizes) the fact-finder's expected utility (loss). Leaving aside technical aspects regarding the nature and feasibility of probability and utility assignment (which are necessary for computing expected values), the contested issue appears to be how to *interpret* decision-theoretic results in practical applications, such as the law. From one point of view, determining the decision(s)[178] with the maximum (minimum) expected utility (loss) is a mere technical operation with some computational burden. However, an additional step is involved until an actual decision is made in practice, and it is legitimate to ask to what extent decision makers are willing to base their decisions on the sole consideration of decision-theoretic advice. This, in turn, raises the broader question of how to understand normative[179] theories, of which decision theory is a prominent example. A widely held view is that, by its nature, a normative account is an imperative for what ought to be done, i.e. what one ought to decide. Presumably, it is this strict interpretation[180] of

---

[176] E.g, Cheng (2013), supra note 164. Note that the scoring rule development presented in §II.B. is based on this idea. According to this scheme the optimal value to report as a probability is the value α that minimizes the expected score.

[177] Allen (2003), supra note 165.

[178] We include the plural because there may be more than one decision with maximum (minimum) expected utility (loss).

[179] In the context here, the term 'normative' does not refer to a legal norm or principle. Instead the meaning is that admitted in the disciplines of judgment and decision-making in applied psychology and artificial intelligence, among others, that rely on philosophical and mathematical arguments to define (logical) standards of evaluation. These so-called normative models are developed and justified independently of observable decision behavior, and make no claim to describe how people actually behave (e.g., Johnathan Baron (2012), The Point of Normative Models in Judgment and Decision Making, 3 Frontiers in Psychology 1–3 (2012); Eric J. Horvitz, John S. Breese & Max Henrion, Decision Theory in Expert Systems and Artificial Intelligence, 2 International Journal of Approximate Reasoning 247–302 (1988)).

[180] Historically, strictly prescriptive interpretations were also promoted for the use of probability theory as a standard for reasoning under uncertainty in the law, a view known a

normative models in the law that provokes objections among legal scholars. The question of how to interpret normative theories is not specific to the law, however, as it has also been a concern for probability and decision theorists, at least since half a century ago, and it is instructive to attend to some of their views. A noteworthy example is the position of de Finetti, whose decisional account was presented in §II.B. Although considered by some as a radical probabilist,[181] he favored a well reflected and balanced view of decision theory. He explicitly recognized the interpretation of a normative theory as a "compendium of mandatory prescriptions"[182] as problematic because it would lead to a suppression of freedom of choice. Recognizing critiques of strict normativism, de Finetti considered it necessary to view normative advice as conditional:

> "(...) I consider the criticisms of aspects of a theory that contain something rigidly normative as justified, while the normative character in the conditional sense must be considered not only acceptable but necessary." [183]

Thus, according to de Finetti, a decision-theoretic result should serve not as a prescription, but as a *conditional advice*, contributing to the entirety of considerations taken into account by the decision maker, prior to making a decision:

> "(...) 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."[184]

'Bayesian imperialism' (e.g., Ward Edwards, Influence Diagrams, Bayesian Imperialism and the *Collins* Case: an Appeal to Reason, 13 Cardozo Law Review 1025–1074 (1991)).

[181] Maria Carla Galavotti (Ed.), Bruno de Finetti, Radical Probabilist (2009). Even de Finetti himself, on occasion, used the word 'radical' (e.g., "the radical nature of certain of my theses", de Finetti, supra note 7, at xvii).

[182] [transl. by the authors; original: "recueil de prescriptions obligatoires"] Bruno de Finetti, Dans quel sens la théorie de la décision est-elle et doit-elle être "normative", Colloques Internationaux du Centre National de la Recherche Scientifique, "La Décision", Paris (25–30 May 1960), 159–169, p. 163.

[183] Id. [translation by the authors; original: "(...) je considère justifiées les critiques à des aspects d'une théorie qui contiendraient quelque chose de rigidement normatif, tandis que le caractère normatif dans le sens conditionnel doit être considéré non seulement acceptable mais nécessaire."]

[184] Id. [translation by the authors; original: "(...) il ne s'agit nullement de prescrire quoi que ce soit, mais seulement de donner à chacun un outil pour lui permettre de mieux analyser et de comparer le pour et le contre de chaque décision selon le point de vue qui lui est propre, avant de choisir."] Similarly, Jim Q. Smith has noted: "The analyst's task is never to tell the DM [decision maker] what to do but to provide frameworks to help her creatively explore her problem and to come to a reasoned decision she herself owns (...)" (Jim Q. Smith, Bayesian Decision Analysis, Principles and Practice (2010), at p. 26).

As we have argued elsewhere,[185] this perspective seems well compatible with the relative plausibility account, according to which participants in the legal process rely on a broad scope of cognitive devices and strategies in their reasoning and decision-making activities.

4. Numeracy and the quest for practicality

Critical readers might ask how a theory can produce advice of any kind if it is based on formulae that involve numerical components for which no guidance is offered regarding how to assign them. Bell, for example, has noted that the formula proposed by Kaplan[186] "does not rationalize any particular standard of proof unless values are assigned to the variables"[187]. Apart from being a truism, this critique of a formula is in fact a critique of the concept of utility on which the formula is built. The assignment of utilities is a notoriously difficult topic, a difficulty we do not deny. Notwithstanding, it is important to be clear as to the exact purpose of the notion of utility: in essence, it is a number that expresses the value *a person assigns* to a particular decision consequence (i.e., outcome).[188] Thus, utility is operationalised only through its application by a person who uses it to express her preferences between decision consequences. Any mechanism of autonomous value assignment would make the concept vacuous, as it would deprive it from the very purpose it is designed for: helping persons to make a value assignment.

It is sometimes suggested that, as a consequence of the difficulty of assigning utilities, the concept should be abandoned altogether. One is free, of course, to disregard the concept of utility, but this would not make the problem disappear. One would still face the inevitable reality that decisions *do have* consequences and that we *do have* varying preferences between those consequences. Any formal approach to the problem would require one to provide a means to capture the value of decision consequences. Utility is one such means. Again, critics may question whether a single number (i.e., utility) can meaningfully capture something so complex as the value of decision consequences in legal applications. Admittedly, a single-valued utility may be considered as a simplistic measure, but how could we ever pretend to employ a more elaborate measure if we cannot even handle a simple

---

[185] Alex Biedermann, Silvia Bozza & Franco Taroni, Normative Decision Analysis in Forensic Science, Artificial Intelligence and Law (forthcoming).

[186] Kaplan (1968), supra note 23, at 1072. Kaplan's formula specifies the threshold probability necessary to warrant a particular legal conclusion as a function of personal preferences among the various possible decision outcomes, expressed in terms of what he called "disutilities or regrets" (at 1079).

[187] Bell (1987), supra note 156, at 560–561.

[188] Utility (a number) expresses the relative strength of preferences, though expressing preferences as an ordering is enough (ordinal utility).

measure?[189]

Although it is worthwhile scrutinizing assignments of particular numerical values, it would be inappropriate to consider 'hard' numerical assignments as a necessary requirement to make decision-theoretic formulae useful.[190] To illustrate this point, consider the well-known result that, in the ordinary civil case, Kaplan's[191] formula leads to the threshold probability of >0.5 when an erroneous finding for either party is assigned the same disutility (e.g., a symmetric 0-*k* loss function).[192] This value, >0.5, is widely considered as the numerical expression of the 'preponderance of the evidence' standard for the typical civil case.[193] This is not the main point, however, since the formula shows more than this. Even without settling for particular numbers, increasing the loss value assigned for the adverse result of one decision, compared to the loss value assigned to the adverse consequence that can potentially result from the alternative decision, will lead to an increase of said threshold probability (above 0.5) – as commonly required in the context of criminal cases. Friedman has concisely summarised these qualitative relationships as follows:

> "Suppose that Option One has far worse consequences if wrong than does Option Two. Then a sensible decision-maker will choose Option One rather than Option Two only if she has a high degree of confidence that Option One rather than Option Two is correct, or, put another way, only if she thinks Option One is far more probable than Option Two."[194]

More informally, the decision-theoretic criterion thus appears to capture the common intuition that 'the more is at stake, the more one shall be sure before one decides', that is the understanding that one's decision is, broadly speaking, a 'weighing' of relative strengths of belief (i.e., the extent to which one is convinced of the truth or otherwise of competing propositions) against

---

[189] Lindley (1985), supra note 47, at 65 ("The truth of the matter is that people dislike simple problems: they like to take refuge in complicated ones where inadequacies of their procedures are difficult to challenge because of the obscurity generated by complication. (...) The reply to the accusation of guessing at probabilities and utilities is simply that if you can't do simple problems, how can you do complicated ones?").

[190] More generally, numeracy is not considered a necessary component of juridical proof for some accounts, such as the relative plausibility account, are explicitly non-numerical.

[191] Kaplan (1968), supra note 23, at 1072.

[192] E.g., Kaye (1999), supra note 147, at 1.

[193] Some scholars consider this assumption mistaken (e.g., Allen & Pardo (2018), supra note 5, and Michael Pardo, The Paradoxes of Legal Proof: A Critical Guide, 99 Boston University Law Review 233–290 (2019)).

[194] Friedman (2018), supra note 9, at 1590

relative losses.[195,196] Clearly, this qualitative relationship can be considered at various levels of detail, as argued by Friedman: "The concept of *magnitude* is essential. (...) we need to have a sense of *how much* worse one error is than the other."[197]

In combination, the above strains of argument emphasize the analytical nature of the decision-theoretic perspective and its role in providing a structured way of thinking about problems, in a conditionally advisory sense, as compared to accounts that focus primarily on providing a descriptive account of decision behaviour, or the functioning of a system.[198] Further, the non-prescriptive interpretation of the formal decision-theoretic perspective avoids the reduction of practical decisions to purely mathematical questions, as critics may invoke. Specifically, it does not intend to delegate decision competence or decision prerogatives to an abstract criterion or a theory. Instead, the purpose of conditional decision-theoretic advice is to question and scrutinise one's consistency relationships (i.e., coherence) among (probabilistic) beliefs and value assignments for decision consequences, in particular adverse outcomes, and make these considerations formally precise, so as to help justify particular lines of reasoning or conclusions.

PART 4: DISCUSSION AND CONCLUSIONS

Past and contemporary debates among legal scholars on the nature, substance and form of evidence evaluation and proof processes in the law show a stark tendency to develop comprehensive theories. The construction of such theories is ambitious, however, and the various participants to the discussion have found their counterparts' accounts deficient in various respects, highlighted often in intense discussions over detailed aspects – provoking equally detailed replies and rebuttals.[199] As argued in a recent work by Professor Michael Pardo,[200] too often, there has been a focus on treating problems in isolation, thus making it difficult to see connections between and broader dimensions of the different concepts.

---

[195] Biedermann et al. (2018), supra note 185.

[196] This is a formal argument in support of the view that purely probabilistic criteria are conceptually insufficient for warranting decisions (see also §3.A.).

[197] Friedman (2018), supra note 9, at 1590.

[198] This does not mean, however, that the formal decision-theoretic account does not offer descriptive potential. As noted above, it is indeed capable of providing a rationale for some common understandings of standards of proof, especially their numerical (i.e., probabilistic) components and the valuation of decision consequences (i.e., their relative adversity).

[199] Some commentators consider that, over the past few decades, the discussions have exhausted themselves and that the various sides merely retreat to (re-)affirming their established positions (e.g., Allen (2018), supra note 134; Friedman (2018), supra note 9).

[200] Michael Pardo (2018), supra note 193.

One area where this difficulty arises is the nature and role of probability in the legal process. It is often said, for example, that the rules of probability map poorly onto the way in which people actually think (when unaided from a cognitive standpoint). Although being an empirical truth,[201] this does not – as we have argued through this paper – impede on the core of probability theory for the latter does not claim to provide a descriptive account. Similarly, no one would find arithmetic operations of addition and subtraction unsuitable simply because some people find them difficult to operate or understand.[202] Thus, critiques should be tailored to what the respective accounts actually purport to do, and proponents of particular accounts – especially probabilists – should be careful not to overinterpret the scope of applicability of their accounts.[203] Better delineating the potential and scope of formal methods of reasoning and decision analysis should help us (i) further the understanding and clarify key aspects of legal proof and decision processes, (ii) prevent unjustified criticisms and (iii) shed more light on ways in which distinct concepts can beneficially interact so as to offer solutions that go beyond what individual accounts may be able to provide in isolation.

This article started from the premise that if particular formal concepts, especially probability and decision theory, have a place within the broad scope of cognitive devices used in applied contexts such as the law, it remains important to address the recurrent criticism to which these concepts are exposed. Despite the fact that this might be perceived as treating individualised and local problems, resolving these issues remains an important preliminary to deploying the full potential of the various concepts.

This article makes three contributions in this regard. The first is a critical review and discussion of recurrent critiques of probability. The second is the presentation of a historically important and logically well founded, but in the law largely unknown, account of probability. It is based on the idea of encouraging individuals to take responsibility for their quantifications of uncertainty, by viewing probability assertion as a decision, rather than a deference to abstract definitions that conflate the definition of probability with its assessment in actual cases, which encounters operational difficulties. The third contribution emphasizes the conceptual relationship between this decisional perspective of probability and other notions and decision points in the legal process. Here our argument focuses on what we call decision

---

[201] E.g., Daniel Kahneman, Paul Slovic & Amos Tversky, Judgement Under Uncertainty: Heuristics and Biases (1982).

[202] But clearly, if you are supposed to receive two payments of 100$, but your bank decides to increase your account only by 180$, it will be in your interest to insist on receiving the full amount of 100$+100$=200$, unless you are willing to lose the difference of 20$. You are free, of course, to ignore advice of formal analyses, but this liberty comes at a price.

[203] Allen (2015), supra note 85.

structures that represent the minimal anatomy of decision problems as they appear to all discussants, irrespective of the position from which they argue. In analysing these decision structures, we invoke elements of classic decision theory, but emphasise the need to interpret them carefully. Here, too, we adopt a historical perspective and rely on foundational literature to argue against overinterpreting results of normative analyses in a rigidly prescriptive sense. We identify such overinterpretation as a further source of controversies over the role of probability and decision theory in the law.

Regarding our first topic of inquiry, probability interpretations in the law, we found it inevitable to review the divide between objectivist and subjectivist viewpoints, both of which face distinct critiques. While we agree with the shortcomings and limitations of objectivist accounts of probability, which demonstrate their illusionary goals and self-defeating character, we insist on considering these complications not as a drawback of probability as such, but only as a deficiency of a particular interpretation of probability. This leaves the subjectivist interpretation of probability as a remaining option,[204] though it is not free of attracting its own critiques. Among the most widely known critique is the sentiment that 'subjective' means allowance of arbitrary and unconstrained assertions, in the sense of 'anything goes', in ignorance of any available evidence. Where probability assertions indeed exhibit these shortages, legal scholars and practitioners are right to expose and denounce them.[205] In such instances, however, their critiques do not address an inadequacy of subjective probability, but a misapprehension of what it really means to understand and apply probability in terms of subjective degrees of belief. We have illustrated this through an analysis and discussion of relevant historical sources that depict an image of subjective probability, based on the notion of *justified assertion*, that is quite different from how it is exposed in standard critiques. While we concede that, in general, probability is a very liberal concept with rather few constraints, apart from requirements of consistency, it is desirable to conceive of a way to encourage honest and well-reflected probability assertions.

This has led us to our second and main topic of inquiry: the understanding of probability assertion as a decision. The decisionalization of probability, as we call it, highlights the idea that asserting probabilities should be assessed carefully, making the best use of all relevant available data, knowledge and

---

[204] Although critical of all interpretations of probability, Professor Allen on at least one occasion has noted that "[none] of the conceptualizations of probability except probability as subjective degrees of belief can function at trial." (Ronald J. Allen, Taming Complexity: Rationality, the Law of Evidence and the Nature of the Legal System, 12 Law, Probability & Risk 99–113 (2013), at 104).

[205] For discussion of this observation in probability applications by forensic scientists, see Biedermann (2017), supra note 52, at 478–479.

information – in a way that can be disclosed and audited. The underlying formal development, based on (proper) scoring rules, is rather well-known in the history of subjective probability, but rarely invoked in legal literature and science other than some areas involving the elicitation of expert opinions and the evaluation of forecasters. The reasons for this are, supposedly, its uncompromising rigour and intricacy. Clearly, thus, our suggestion here is not to introduce probability elicitation based on scoring rules in practical proceedings. We have invoked scoring rules solely for the purpose of showing that there is a way to formalise the intuition of considering probability assertion as a decision, with the interesting result that one should not – under a proper scoring rule – *decide* to report a distorted probability that does not correspond to one's actual belief (e.g., rounding off a small probability to zero). Notwithstanding, scientific (e.g., forensic) experts may be expected to know about and understand the related theory.[206] The practical avenues for this perspective remain vague, however, since there is no structured approach to probability in science; there remains a strong tradition of educating scientists in mainstream accounts[207] of probability, without informing them enough about the conceptual limitations of these approaches, let alone exposing them to critical methodological enquiry about foundational concepts – areas often considered too far from practical consideration and, at best, to be left to the scrutiny of philosophers of science or statisticians with special interests.[208]

A further reason why we have introduced the decisional perspective to probability is that it is merely a particular instance of the application of classic decision theory that reveals structural analogies with respect to other stages of juridical evidential assessment. Having clarified how it structures the logic of probability assignment in terms of fundamental ingredients such as decisions, uncertain states of nature, and decision consequences, we have argued in part 3 that the logic of the decisional perspective is not fundamentally different from typical legal decision problems, such as decisions about ultimate issues at trial. While we recognise that the application of formal decision-theoretic analyses in the law is a controversial topic in legal scholarship, we identify one root of this controversy in disagreements over the deeper point of how to understand results of formal, conceptually normative and analytical accounts, as opposed to descriptive theories of how (legal) systems actually work and of observable decision behaviour of people operating in such systems. In our view, the question is not whether or not litigants or legal processes in general are or ought to be

---

[206] Biedermann (2013), supra note 118.

[207] So-called "*cookbooks* of the standard literature" (Piccinato (1986), supra note 137, at 13) or "Adhockeries" (de Finetti (2017), supra note 7, at 11).

[208] Lad (1996), supra note 88.

acting in agreement with the results of formal decision-theoretic analyses. The theory (of decision) is more general: the point is that it provides a framework that allows one to subject *any* decision – whether it is one's own decision or that of others – to a formal analysis *independently* of whether or not one actually endorses the theory. The result of such analyses will not pre-empt or prescribe any particular decision, but represent advice that will add to the entire scope of considerations upon which a practical decision is based. As an aside, these analyses expose traditional discourses that treat standards of decision coarsely as purely probabilistic thresholds, as conceptually incomplete because they do not explicitly address the valuation of uncertain decision consequences, some of which may be adverse.

In conclusion, we find that the decisional perspective supports a logically sound and analytically rigorous examination of underlying threads and commonalities of key topics – i.e., probability assertion, plausibility assessment and legal conclusions about ultimate issues – that otherwise tend to be dealt with in literature either individually or in opposing ways. A better appreciation of the various concepts' respective scope of capacities and limitations should help us to use them in more defensible and complementary ways. In particular, the competing accounts' descriptive, normative and interpretative properties require careful attention, as we have shown, to avoid misconceived critiques and overinterpretation of theories beyond their domain of competence.