



# Decision Theory, Relative Plausibility and the Criminal Standard of Proof

Alex Biedermann<sup>1</sup> · David Caruso<sup>2</sup> · Kyriakos N. Kotsoglou<sup>3</sup>

Published online: 19 February 2020  
© Springer Nature B.V. 2020

## Abstract

The evolution of the understanding of evidence-based proof and decision processes in the law, especially criminal law, and standards of proof in this area, has a long-standing and controversial history. Competing accounts cause the legal scholarship to engage in critical and thoughtful exchanges. Some of the divergent views reflect different methodological perspectives similarly recognized in other fields, such as applied psychology and economy, and the broader interdisciplinary research fields of judgment and decision-making, system analysis and decision science. One such methodological perspective asserts that accounts of juridical proof should provide a description and explanation of how the legal system actually works as a whole. Other—more mathematical and analytical accounts—concentrate on how, ideally, legal decision-making under uncertainty ought to be made in order to be considered sensible. This paper focuses on the relative plausibility (RP) account advocated by Professors Allen and Pardo as an example of the former perspective. Its logical structure and argumentative implications are analysed using elements of decision theory, which is the prime representative of the latter, more mathematical approach to legal proof. Using formal diagrammatic schemes to depict the structural relationships between the core elements of the two accounts, it is demonstrated in what sense they can be considered logically related and congruent. The demonstration shows that the principal disagreements among the proponents of the two examined theories derive from differences in (1) the criteria used to judge the adequacy of competing accounts of legal decision-making, and (2) the level of formalization of the bases of decisions in each candidate account. This structural analysis supports the view that adherence to one or the other of the examined perspectives does not imply a contradiction, but reflects the coverage of different aspects of the same overall decision architecture. Using decision-theoretic notions, our analyses also provide a way to explain RP decisions through an explicit criterion, thus providing a reply

---

The authors thank Ronald Allen of Northwestern University Pritzker School of Law and Paolo Garbolino of Iuav Università di Venezia for their helpful comments and suggestions. Alex Biedermann gratefully acknowledges the support of the Swiss National Science Foundation through grant No. BSSG10\_155809.

---

Extended author information available on the last page of the article

to the recurrent critique that RP theory lacks specific means to justify its decisional framework.

**Keywords** Standard of proof · Decision theory · Relative plausibility · Probability · Influence diagrams

*Our approach does not eliminate the need for you to make judgments and to express preferences; anyone who so claims is a charlatan.*  
(Irving H. LaValle, *Fundamentals of Decision Analysis*, 1978, at 13, *italics as in original*).

## 1 Introduction

### 1.1 The Relative Plausibility Account of the Criminal Standard of Proof

In a series of papers, Professors Ronald J. Allen and Michael Pardo have developed and defended an explanatory account of juridical proof based on the notion of ‘relative plausibility’, which we shall refer to as the theory of “RP” as a short-hand.<sup>1</sup> Unlike probabilistic conceptualizations, their theory does not rely on any particular numerical threshold, such as the commonly featured 0.5 threshold for the typical civil case, or any other particular numerical value for more austere standards of proof. Instead, RP is essentially comparative in nature and concentrates on competing explanations advanced by parties. In this context, the notion of explanation is understood in at least two different dimensions. In one dimension, standards of decision for individual cases are conceptualized as inquiries into the extent to which a given party’s story can account for—or so to say ‘explain’—the evidence in a case and the objects of the party in the litigation. In the second dimension, RP is considered explanatory—in a descriptive sense—for what actually happens at trial, that is how the law structures the legal process as whole<sup>2</sup> and how litigants behave.<sup>3</sup>

These two dimensions of explanation in the RP account are apt to the civil contest. In civil disputes, contesting parties each seek to explain the evidence according

<sup>1</sup> E.g., Ronald J. Allen & Michael S. Pardo, Relative Plausibility and its Critics, 23 *The International Journal of Evidence & Proof* 5–59 (2018).

<sup>2</sup> Ronald J. Allen & Michael S. Pardo, Clarifying Relative Plausibility: A Rejoinder, 23 *The International Journal of Evidence & Proof* 205–217 (2019) (“relative plausibility is not only about jury decision-making and the decision rules at trial. Its scope is much broader. It is about the entire process of proof” at 207), Allen & Pardo, *supra* note 1 (“What is being observed is the entire litigation process. (...) what is the best explanation of the data, where “the data” are observations of how the American legal system structures proof at trial” at 7).

<sup>3</sup> Michael S. Pardo & Ronald J. Allen, Juridical Proof and the Best Explanation, 27 *Law and Philosophy* 223–268 (2008) (“Understanding the standards in terms of competing explanations more accurately describes what occurs at trial, [and] is consistent with the [sic] our best understanding of the reasoning processes of jurors” at 261), Allen & Pardo, *supra* note 2 (“The primary message of relative plausibility is that from the beginning to end the legal system pushes the parties to provide competing explanations, and these explanations structure the decision that is subsequently made (even if the decision is based on an explanation not advanced by the parties).” at 208).

to their objects. The legal paradigm for resolution of this contest places the explanations of each party in *relative* comparison apropos which is (more) plausible: this is the product of the preponderance standard. Suppose a standard civil claim. A lends B 500 USD, but later claims that B has defaulted on the loan and sues for repayment. B denies this allegation and argues that he never actually received the money from A. For such a case the RP account amounts to, in essence, the fact-finder assessing whether the plaintiff or the defendant produces a better<sup>4</sup> account, given all the evidence presented.

The criminal case involves somewhat wider considerations. Allen and Pardo describe them as follows:

In criminal cases, under the ‘beyond a reasonable doubt’ standard, the prosecution must do more than present a *better* explanation than the defense (or the best available explanation): fact-finders convict only when the prosecution’s explanation (which includes all of the legal elements) is plausible, given the evidence, and there is *no* plausible defense explanation.<sup>5</sup>

The RP account thus focuses on examining whether explanatory thresholds are met. This involves a staged scheme of plausibility assessment, followed by particular legal conclusions of conviction or acquittal. According to this scheme, the first issue is whether the prosecution has a plausible story, capable to warrant a guilty verdict for if there is no such account, the defendant is to be acquitted. If such a plausible account is provided by the prosecution, then, second, the fact-finder has to assess whether the defense offers a plausible story. If a plausible defense account is provided, then the defendant merits an acquittal,<sup>6</sup> even if the defense account is considered less plausible than the one offered by the prosecution.

Allen and Pardo describe the criminal case in a manner that places the prosecution and defence in positions *relative* to each other. To be clear, central to their view is that there is *no* external measure against which each side is assessed. This is in contrast to the more conventional understanding that regards the legal requirement of beyond reasonable doubt for criminal litigation, unlike that in civil, to change the respective positions of the parties, not as relative to each other, but as relative to a legal standard. That is, the prosecution’s account is assessed relative to what is embodied in the notion of beyond reasonable doubt; the account of the defence (whether through adducing evidence or challenging the prosecution) is relative to what is encompassed within the notion of a reasonable doubt. While the concept of RP is completely different from this conventional explanation of criminal litigation, RP as a model is attractive to explain empirical aspects of civil disputes because it can relate the accounts of the parties to each other as a result of the civil standard requiring nothing more than an adjudgment of the relative merits of the competing

<sup>4</sup> The critical reader might immediately invoke the question of how one is to understand ‘better’. This question will be discussed later in § 2.1.

<sup>5</sup> Allen & Pardo, *supra* note 1, at 16.

<sup>6</sup> *Id.* (“any plausible explanation consistent with innocence (one that has not been disproven or eliminated) is sufficient to raise a reasonable doubt” at 27).

accounts. It is, of course, reasonable to consider that courts in criminal cases may relate the accounts of prosecution and defence to each other. That is, if the prosecution presents a plausible account which is not plausibly infringed by the defence, the fact finder will convict. In this sense, RP offers a model to explain the objects of litigants (the first dimension) and the objects of trial (the second dimension).

The application of the model to the criminal trial may be difficult, however, for mindsets strictly adhering to the view that the plausibility of the account proffered (by prosecution or defence) is relative to the standard of proof required. The result is a dearth of revelation regarding how decision-makers resolve criminal disputes because the content and concept of reasonable doubt, and what takes us beyond that (for conviction), remains opaque. But, Allen and Pardo do not purport to resolve such matters through their model. Their point is to emphasize that the explicitly non-numerical nature of their account is an advantage over the probabilistic account. They argue that controversies over how to *exactly* define particular standards of proof (i.e., numerically and in terms of fixed ‘probability points’), such as the beyond a reasonable doubt or the “notoriously obscure clear-and-convincing-evidence-standard”,<sup>7</sup> are naturally and effectively avoided. Clearly, by focusing the attention on what practitioners can and are said to already do, i.e. operating in qualitative and comparative ways, the RP account should gain attractiveness. But critical commentators regard this feature as a weakness. Professor Nance, for example, questions “whether this [=plausibility] advances our understanding discernibly”.<sup>8</sup> He notes that the RP account offers no clear answer to the issue of “*how strong* must one’s doubt be before it becomes a *reasonable* one”,<sup>9</sup> and that “[t]his question is not illuminated by being told a doubt is strong enough to be reasonable when the pro-defense explanation is ‘plausible’.”<sup>10</sup> To some extent, it thus seems as though we face a paradoxical impasse: a controversial aspect—the question of quantitative thresholds—that RP dissolves, is, in turn, invoked as a feature that it lacks. This impasse reveals an area for future research, especially regarding “possibilities for formalizing aspects of relative plausibility (...) [and] compatibilities between the explanatory and probabilistic perspectives”.<sup>11</sup>

---

<sup>7</sup> Id., at 29.

<sup>8</sup> Dale A. Nance, The Limitations of Relative Plausibility Theory, 23 The International Journal of Evidence & Proof 154–160 (2019), at 157.

<sup>9</sup> Id., emphasis as in original.

<sup>10</sup> Id.

<sup>11</sup> Allen & Pardo, *supra* note 2, at 215. A number of evidence scholars have already questioned the intelligibility of an all-or-nothing choice between atomistic and holistic approaches. For more discussion see e.g. William Twining, *Hot Air in the Redwoods, A Sequel to The Wind in the Willows*, 86 Michigan Law Review 1523–1547 (1988), at 1543.

## 1.2 Contents and Structure of this Paper

In this paper, we examine the RP concept using elements of decision theory. Our analysis will rely on the idea of interpreting plausibility assessments as *decisions* that are to be made in the light of uncertainty that inevitably attends all real-world proceedings. We believe that regarding RP assertions as decisions is descriptively valid and minimally constraining. As a methodological starting point, it also looks compatible with the way in which the developers of the RP theory describe their account:

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.<sup>12</sup> [emphasis added]

The choice of a decisional perspective as a starting point is intricate, however, because formal decision analyses are regarded critically in this context. For example, Allen & Pardo have noted:

(...) we doubt that there will be a well-organized science (or philosophical thought) of how people *decide* one explanation is better than another, even though that is precisely what people do in a virtually infinite number of ways.<sup>13</sup> [emphasis added]

We thus seek to be clear about the sense in which we intend to pursue the analysis of RP from a decision-analytic perspective. Our analysis will not claim to provide a descriptive account of how people decide, or supposedly decide. Instead, we will seek to capture (1) the inevitable and fundamental structural elements of the ‘problem’ of *deciding* about RP as faced by anybody who intends to apply the plausibility scheme, and (2) the logical relationships among the elements thus identified.

Likewise, the decision-analytic account that we provide does not purport to *prescribe* how one ought to decide, neither in general nor in particular cases. Rather, our account aims to clarify what we shall call the anatomy of the decisional task. It seeks to help decision-makers conceptualize and structure their assessment task (of RP) in a transparent and logically defensible way. We anticipate that this analysis will show that there is a neat and non-conflictual relationship between RP *decisions* and the formal decision-theoretic account of legal decision-making as described in existing literature.<sup>14</sup>

Our paper is structured as follows. Section 2 will introduce elements of classic decision theory, including standard graphical methods for decision analysis (i.e.,

<sup>12</sup> Allen & Pardo, *supra* note 1, at 6. Similarly, at 34, the same authors note that “the fact-finder is essentially asked to *decide* which [source of causation] is more plausible”.

<sup>13</sup> Allen & Pardo, *supra* note 2, at 210.

<sup>14</sup> The first formal decision-theoretic account of legal proof is widely attributed to John Kaplan, *Decision Theory and the Factfinding Process*, 20 *Stanford Law Review* 1065–1092 (1968).

decision trees and influence diagrams). In Section 3, we will use these concepts for analysing the scheme for making RP judgments in criminal cases.<sup>15</sup> Critical discussion and conclusions will be presented in Section 4.

## 2 Elements of Classic Decision Theory and Their Application in the Law

### 2.1 Formal Development

Over the decades since its publication in 1968, Kaplan's foundational paper<sup>16</sup> on decision theory in the law has been mentioned in court cases<sup>17</sup> and provided grounds for widespread discussion among legal scholars.<sup>18</sup> Kaplan's main result, also sometimes referred to as the "Kaplan formula",<sup>19</sup> is also featured in contemporary textbooks.<sup>20</sup>

In a nutshell, a decision-theoretic analysis relies on the following notions: decisions (or courses of action), states of nature (i.e., what is actually the case) and decision consequences; we will further explain these terms in due course. The analysis also involves probabilities for uncertain states of nature, expressions of preferences (measured by utilities or losses) among possible decision consequences and expected values (i.e., expected utilities/losses) serving as expressions of the desirability of decisions. In combination, these elements represent an elaborate technical apparatus which naturally offers many angles of attack.<sup>21</sup> For the time being, we shall ask our readers to bear with us and consider a basic illustration of how the above notions can be interpreted in legal contexts, keeping in mind that, as noted by Professor David Kaye, "no mathematical result is self-applying, and additional

<sup>15</sup> For a decision-theoretic analysis of relative plausibility decisions in civil cases, see Alex Biedermann & Joëlle Vuille, *The Decisional Nature of Probability and Plausibility Assessments in Juridical Evidence and Proof*, 16 *International Commentary on Evidence* 1–30 (2018).

<sup>16</sup> *Supra* note 14.

<sup>17</sup> E.g., *In re Winship*, 397 U.S. 358 (1970), at 370 (Harlan, J., concurring); *United States v. Parr*, 516 F.2d 458 (5th Cir. 1975), at 464; *California ex rel. Cooper v. Mitchell Brothers' Santa Ana Theater*, 454 U.S. 90 (1981), at 93.

<sup>18</sup> E.g., Richard S. Bell, *Decision Theory and Due Process: A Critique of the Supreme Court's Law-making for Burdens of Proof*, 78 *The Journal of Criminal Law & Criminology* 557–585 (1987); Bernard Grofman, *Mathematical Models of Juror and Jury Decision-Making: The State of the Art*, in: Bruce D. Sales (Ed.), *The Trial Process* 305–351 (1981); David H. Kaye, *Clarifying the Burden of Persuasion: What Bayesian Decision Rules Do and Do Not Do*, 3 *The International Journal of Evidence & Proof* 1–28 (1999); Richard O. Lempert, *Modeling Relevance*, 75 *Michigan Law Review* 1021–1057 (1977).

<sup>19</sup> Bell (1987) *supra* note 18, at 558.

<sup>20</sup> E.g., Dale A. Nance, *The Burdens of Proof: Discriminatory Power, Weight of Evidence, and Tenacity of Belief*, 2016, at 23; Richard D. Friedman, *The Elements of Evidence* (4th Ed.), 2017, at 592 (presenting a more general formula).

<sup>21</sup> E.g., Ronald J. Allen, *The Error of Expected Loss Minimization*, 2 *Law, Probability & Risk* 1–7 (2003); Allen & Pardo, *supra* note 1.

argument is necessary to bridge the gap from a general mathematical truth to a substantive application—in law as in any other domain.”<sup>22</sup>

Start by considering two competing assertions, one provided by the prosecution, denoted  $H_p$ , and one provided by the defense, denoted  $H_d$ . These two assertions, also sometimes called propositions, refer to a contested event of interest. Simply put, they stand for the possible states of nature: contested versions of the way in which the world is or has been. Because there is uncertainty about which state of nature (or, version of the contested event) is true, the trier of fact holds probabilities for  $H_p$  and  $H_d$ , denoted  $\Pr(H_p)$  and  $\Pr(H_d)$ , respectively.<sup>23</sup> Next, denote by  $d_p$  (conviction) and  $d_d$  (acquittal) the two<sup>24</sup> possible courses of action (or, decision), one of which must be taken.<sup>25</sup> We write  $C_{ij}$  for the consequence that follows deciding  $d_i$  when in fact the proposition  $H_j$  is true. Clearly, for an omniscient observer, a decision is either accurate or erroneous: an accurate finding is one for the prosecution (defense),  $d_p$  ( $d_d$ ), when in fact  $H_p$  ( $H_d$ ) holds, otherwise it is erroneous. Most naturally, an accurate finding is preferred to an erroneous finding. Often, preferences among decision consequences are expressed (or, measured) using the concepts of utility or loss. For the purpose of the current analysis, we will use the notion of loss, written  $L(C_{ij})$ . Loss is a positive number expressing the extent to which the decision-maker considers a given adverse consequence  $C_{ij}$  undesirable; we will set the maximum loss that may be incurred to 1. Accurate findings are assigned a zero loss because it is considered that nothing is lost by making a correct determination. Such a loss function is also known in literature as a 0–1 loss function.

At the time of making a decision in a criminal case, and at all other points in time (future or past), it is not known with certainty which state of nature actually holds. For whatever side one will find, there is an inevitable possibility for one’s decision to lead to an erroneous consequence, depending on the probability of  $H_p$  and  $H_d$ , respectively. Stated otherwise, it is impossible to control the actual loss incurred. At best, one may consider the *expected losses* associated with the available decisions. The expected loss, written  $EL(d_i)$ , is obtained as follows:

$$EL(d_i) = L(C_{ip}) \Pr(H_p) + L(C_{id}) \Pr(H_d). \quad (1)$$

Briefly stated, the expected loss of a given decision is the sum of the actual losses of each decision consequence, ‘weighted’ by the respective probability of their occurrence.

<sup>22</sup> Kaye (1999), *supra* note 18, at 27.

<sup>23</sup> Note that for shortness of notation, we do not include the conditioning on background information  $I$  and the entirety of the evidence  $E$  available to the decision-maker at the time when the decision needs to be made. A more complete notation for the probability of a proposition  $H$  would be  $\Pr(H|I, E)$  where ‘ $I$ ’ denotes ‘conditioned on’.

<sup>24</sup> Decision-theoretic analyses may be extended to more than two decisions. Also, there may be more than two propositions.

<sup>25</sup> Technically, though, fact-finders do not ‘find’ for the defendant in case of an acquittal in criminal adjudication, for the defendant is presumed innocent. The defendant’s default status is merely preserved.

Consider, for example, the situation in which one decides for the prosecution by convicting,  $d_p$ . The expected loss is  $EL(d_p) = L(C_{pp})Pr(H_p) + L(C_{pd})Pr(H_d)$ , where  $C_{pp}$  is a correct finding for the prosecution, and  $C_{pd}$  is an erroneous finding (i.e., wrongful conviction). It is immediately seen, however, that the loss assignment  $L(C_{pp}) = 0$  reduces the expected loss of decision  $d_p$  to  $EL(d_p) = L(C_{pd})Pr(H_d)$ . Similarly, the expected loss for the decision  $d_d$  (acquittal) is given by  $EL(d_d) = L(C_{dp})Pr(H_p) + L(C_{dd})Pr(H_d)$ . A correct acquittal  $C_{dd}$  has zero loss, which reduces the expected loss for decision  $d_d$  to  $EL(d_d) = L(C_{dp})Pr(H_p)$ .

The formalization introduced so far provides a way to assess the relative worth or merit of the competing decisions, but it does not yet direct one to a particular decision. It is at this point that decision-theory involves an additional notion, i.e. the notion of minimizing expected loss, used as a decision criterion. This criterion says to choose the decision that has the minimum expected loss.

The reader can easily see that one cannot immediately say which decision has the minimum expected loss. This depends on the probabilities and assigned losses. Thus, let us examine in further detail the conditions under which the expected loss of, for example, decision  $d_p$  is smaller ( $<$ ) than the expected loss of  $d_d$ , thus rendering  $d_p$  preferable to  $d_d$ . This can be done by considering the following transformations:

$$\begin{aligned} EL(d_p) &< EL(d_d) \\ L(C_{pd}) Pr(H_d) &< L(C_{dp}) Pr(H_p) \\ Pr(H_p)/Pr(H_d) &> L(C_{pd})/L(C_{dp}) \end{aligned} \quad (2)$$

Equation (2) is the decision-theoretic criterion that specifies the conditions under which finding for the prosecution is the better<sup>26</sup> decision than an acquittal. In essence, Equation (2) amounts to a comparison between the odds in favor of  $H_p$  against  $H_d$ , shown on the left-hand side, and the relative losses associated with erroneous findings shown on the right-hand side (i.e., the loss of a wrongful conviction  $L(C_{pd})$  and the loss of an erroneous acquittal  $L(C_{dp})$ ).<sup>27</sup>

Some interesting properties of Equation (2) are well known among legal scholars. Suppose, for example, that wrongly deciding for either side is considered equally undesirable. This is commonly assumed for the standard civil case. Thus,  $L(C_{pd}) = L(C_{dp})$  and Equation (2) is satisfied whenever  $Pr(H_p) > Pr(H_d)$ , i.e. when  $Pr(H_p) > 0.5$ .<sup>28</sup> This limiting value is commonly associated with the preponderance of the evidence standard. It is also readily seen that this limiting value increases, whenever the ratio  $L(C_{pd})/L(C_{dp})$  increases, thus accounting for higher

<sup>26</sup> In this context, ‘better’ means having the smaller expected loss.

<sup>27</sup> Equation (2) can also be found, in similar form, in statistical literature, e.g., James O. Berger, *Statistical Decision Theory and Bayesian Analysis* (2nd Ed.) 1985, at 164; José M. Bernardo & Adrian F. M. Smith, *Bayesian Theory* (2nd Ed.) 2000, at 391; Giovanni Parmigiani, *Modeling in Medical Decision Making. A Bayesian Approach*, 2002, at 87; Giovanni Parmigiani & Lurdes Inoue, *Decision Theory: Principles and Approaches*, 2009, at 139.

<sup>28</sup> See, e.g., Kaye (1999), *supra* note 18, at 1, for a statement of the same result, but using utilities instead of losses.



standards of proof, as required in criminal cases. Table 1 summarizes some illustrative examples.<sup>29</sup>

As an aside, note that because Equation (2) “works both ways”,<sup>30</sup> it is also possible to consider the limiting value  $\Pr(H_p)$  first, i.e. the standard of proof, and then work out the corresponding loss ratio. This amounts to reading Table 1 from right to left.

## 2.2 The Notion of Error

An issue that is often raised in connection with verdicts is the notion of error, in particular notions such as the risk of error, the allocation of the risk of error (to parties) and the regulation and monitoring of errors (e.g., in terms of their total number). There are two distinct but related perspectives through which the notion of error is commonly discussed. In one perspective, the term ‘errors’—plural—refers to the number (or, proportion) of cases in which erroneous verdicts are issued. This may be of interest when focusing on the functioning of a legal system as a whole. Errors can be of two types: convicting an innocent defendant (also sometimes referred to as ‘false alarm’, false positive or type 1 error), and wrongly acquitting a defendant (type 2 error).<sup>31</sup> The number or proportion of such erroneous findings *across multiple distinct* cases, though real, is not empirically tractable.<sup>32</sup> In another perspective, the focus is on the *given case at hand*. Here, at most one error can be incurred, either an erroneous finding for the prosecution or an erroneous acquittal. This much is rather uncontroversial.

The distinction between the two perspectives is important, however, because statements about the latter are prone to be carried over inappropriately to the former, especially regarding the notions of probability and expected loss introduced in the previous section. We thus emphasise that the notion of probability is understood here as the decision-maker’s assessment that a given party’s version for a particular case at hand is true (denoted,  $\Pr(H)$ ). In turn, the notion of expected loss, computed on the basis of such a case-specific probability, expresses only—as the term literally conveys—the decision-maker’s expectation of loss associated with a particular decision. This does not amount to a statement about the *actual* number (or proportion) of

<sup>29</sup> See Terry Connolly, Decision Theory, Reasonable Doubt, and the Utility of Erroneous Acquittals, 11 Law and Human Behavior 101–112 (1987) for a detailed empirical investigation and discussion of value assignments for decision consequences, including unusual hypothetical situations such as the assignment of a higher utility to an erroneous acquittal than to a correct acquittal.

<sup>30</sup> Bell (1987), *supra* note 18, at 561.

<sup>31</sup> This terminology is based on considering the defense’s case as the null hypothesis and defining a type I error the false rejection of a null hypothesis (e.g., Grofman (1981), *supra* note 18, at 308, and Edward K. Cheng, Reconceptualizing the Burden of Proof, 122 The Yale Law Journal 1254–1279 (2013), at 1260).

<sup>32</sup> Paul Roberts & Adrian Zuckerman, Criminal Evidence, 2nd. Ed., 2010, at 226.

**Table 1** Illustration of the minimum probability (column 4) necessary to ensure that the expected loss for deciding for the prosecution,  $d_p$ , is smaller than the expected loss for deciding  $d_d$  (acquittal), given different loss ratios (column 3)

$L(C_{pd})$	$L(C_{dp})$	Loss ratio	Limiting value $\Pr(H_p)$
1	1	1	0.5
1	0.5	2	0.66
1	0.1	10	0.91
1	0.01	100	0.99
1	0.001	1000	0.999

A loss ratio  $x > 1$  means that the loss incurred by an erroneous decision  $d_p$  (column 1) is  $x$  times greater than the loss incurred by an erroneous decision  $d_d$  (column 2). We assume that an erroneous finding for the prosecution, i.e. a wrongful conviction,  $C_{pd}$ , is the overall worst consequence and thus assigned the loss 1

errors, and their minimisation, across multiple distinct cases.<sup>33</sup> Rather, the expected loss minimisation decision criterion (or, rule) entails the following aspects:

- An error is an adverse decision consequence  $C_{ij}$  that results from deciding  $d_i$  when in fact the proposition  $H_j$  is true, for  $i, j = \{p, d\}$  as defined in §2.1, and  $i \neq j$ .
- What is being targeted is not, however, the possible errors as such, but the characterisation of their undesirability using the notion of loss,  $L(C_{ij})$ . What is more, the focus is not on the *actual* loss, associated with a given error (i.e., an adverse decision consequence), but only the *expected* loss.

Thus, the loss assigned to a particular decision consequence,  $L(C_{ij})$ , is to be distinguished from the *expected* loss associated with a particular decision,  $EL(d_i)$ : the crucial distinction here—and a cornerstone of decision analysis<sup>34</sup>—is the distinction between decisions and decision consequences (or, decision outcomes). Let us recall that because there is uncertainty about the *actual* outcome of a particular decision  $d_i$ , one cannot consider the *actual* loss, but only expected loss, given the sum of the probability-discounted losses assigned to each decision consequence, including both accurate *and* erroneous consequences.<sup>35</sup> So, each decision has its own expected loss which provides a criterion for their comparison. Minimising expected loss then

<sup>33</sup> For a detailed analysis and discussion see Michael L. DeKay, The Difference Between Blackstone-Like Error Ratios and Probabilistic Standards of Proof, 21 Law & Social Inquiry 95–132 (1996), and previously Michael O. Finkelstein, Quantitative Methods in Law: Studies in the Application of Mathematical Probability and Statistics to Legal Problems 65–78 (1978).

<sup>34</sup> As noted, e.g., by Ronald A. Howard (1 Bulletin of the American Mathematical Society 784–787 (1979), at 786) in his review of LaValle (Fundamentals of Decision Analysis, 1978), and in Ronald A. Howard, Decision Analysis in Systems Engineering, in: The Principles and Applications of Decision Analysis, Vol. 1: General Collection, Ronald A. Howard & James E. Matheson (Eds.), 59–93 (1983): “We all want good outcomes. (...) Everyone wants a good rather than a bad (...) – the question is how do we get there. The only thing you can control is the decision and how you go about making that decision. That is the key” (at 92–93).

<sup>35</sup> See also Eq. (1).

means to select the decision that presents, in the view of the decision-maker, the minimum expected loss.

For the remaining parts of this paper, it is therefore important not to equate the limiting probabilities related to particular loss ratios for the individual case through Eq. 2, as summarised in Table 1, with error ratios across many cases. For example, applying a threshold probability of 0.5 does not mean that, across many cases, the ratio of false convictions to false acquittals *will be* 1 to 1. Instead, the loss ratio 1 expresses that, *in the case at hand*, the decision-maker is equally concerned about each type of error that may be incurred.

Statements regarding the number of correct and erroneous verdicts, and hence error ratios, would require knowledge about the proportion of meritorious cases that arrive at trial. Further, assumptions would need to be made about the distributions of the decision-makers' probabilities regarding the parties' versions at the end of each trial, among the two types of cases arriving at trial (i.e., meritorious and non-meritorious cases). Legal literature presents several discussions of these matters, but it is generally acknowledged that they are essentially hypothetical,<sup>36</sup> and thus not further pursued here. We will also take no position regarding the definition of social policy goals, e.g. whether they should be based on considerations of actual errors across multiple cases<sup>37</sup> or on value judgments for case-specific outcomes.

### 2.3 Decision Trees and Influence Diagrams

For readers who are less at ease with formulaic developments as presented in §2.1 we now introduce two graphical concepts, decision trees and influence diagrams. These concepts provide a rigorous graphical mapping of the various components involved in the formal analysis and the relationships among those components. Besides decisions and states of nature, these graphical models incorporate probabilities, losses and expected losses in clear and transparent ways, thus helping to avoid misunderstandings. We insist that, from a decision-theoretic point of view, nothing new is introduced. We are only providing a graphical translation of the decision-analytic apparatus presented so far in exclusively formulaic terms. Readers well acquainted with decision trees and influence diagrams may skip this section. We shall look at decision trees first because influence diagrams are a more compact representation of the same concepts.<sup>38</sup>

<sup>36</sup> For example, Kaye (1999), supra note 18, insists on that the distributions of decision-makers' probabilities across the two types of cases, meritorious and non-meritorious, are "fantasies" (at 24).

<sup>37</sup> As noted above, this is an unknowable factor.

<sup>38</sup> Ronald A. Howard, *From Influence to Relevance to Knowledge*, in: Robert M. Oliver & James Q. Smith (Eds.), *Influence Diagrams, Belief Nets and Decision Analysis*, 3–23 (1990); Jim Q. Smith, *Decision Analysis: A Bayesian Approach*, 1988, at 64.

### 2.3.1 Decision Trees

Decision trees have a long and well documented history in the fields of statistics and applied areas such as the analysis of business decisions.<sup>39</sup> They are also of interest to lawyers who seek to analyse questions arising in professional practice, in particular during the management of cases for clients (e.g., strategic counselling in procedures for filing lawsuits).<sup>40</sup> Here we provide a sketch of decision trees using the simple decision example from §2.1, involving two competing assertions  $H_p$  and  $H_d$  provided by the prosecution and defense, respectively, and two possible courses of action,  $d_p$  (conviction) and  $d_d$  (acquittal).

Figure 1 presents the decision tree for this example.<sup>41</sup> A decision tree is charted horizontally from left, where the trunk (represented by a square) is, to right. The branches that grow out of the trunk correspond to the courses of action available to the decision-maker. Here, there are two branches, representing the two decisions  $d_p$  and  $d_d$ .<sup>42</sup> When pursuing a branch, one arrives at a chance node (also called chance point), presented by a circle. The branches that emanate out of these nodes represent the possible states of nature, or events, about which the decision-maker is uncertain at the time when a decision needs to be made. In the example here, there are two possible states of nature, denoted  $H_p$  and  $H_d$ ,

<sup>39</sup> E.g., John Aitchison, *Choice Against Chance: An Introduction to Statistical Decision Theory*, 1970, at 202–213; Rex Brown, *Rational Choice and Judgment: Decision Analysis for the Decider*, 2005, at 62–64; Simon French, *Readings in Decision Analysis*, 1989, at 27–29; Simon French, John Maule & Nadia Papamichail, *Decision Behaviour, Analysis and Support*, 2009, at 13–21; C. Jackson Grayson, *Decisions Under Uncertainty: Drilling Decisions by Oil and Gas Operators*, 1960, at 323–336 (using the term “Information flow diagram”); Dennis V. Lindley, *Making Decisions*, 1971, at 140–163; James E. Matheson & Ronald A. Howard, *An Introduction to Decision Analysis*, in: Howard & Matheson (1983), supra note 34, 17–55, at 47–51; George E. Monahan, *Management Decision Making: Spreadsheet Modeling, Analysis, and Application*, 2000, at 451–525; Parmigiani & Inoue (2009), supra note 27, at 126–131; Howard Raiffa, *Decision Analysis, Introductory Lectures on Choices under Uncertainty*, 1968; Howard Raiffa, *Decision Analysis: A Personal Account of How it Got Started and Evolved*, 50 *Operations Research* (2000), at 179–185; Jim Q. Smith (1988), supra note 38, at 10–22; Michael D. Resnik, *Choices: An Introduction to Decision Theory*, 1990, at 17–19; Robert Schlaifer, *Analysis of Decisions Under Uncertainty*, 1969 (using the term “decision diagram”, at 37–38); Howard Thomas, *Decision Theory and the Manager*, 1972, at 43–75; Stephen R. Watson & Dennis M. Buede, *Decision Synthesis: The Principles and Practice of Decision Analysis*, 1987, at 36; Robert L. Winkler, *An Introduction to Bayesian Inference and Decision*, 1972, at 219–295; Detlof von Winterfeldt & Ward Edwards, *Decision Analysis and Behavioral Research*, 1986, at 63–89.

<sup>40</sup> E.g., Paul Brest & Linda Hamilton Krieger, *Problem Solving, Decision Making, and Professional Judgment, A Guide for Lawyers and Policymakers*, 2010, at 462–473; David P. Hoffer, *Decision Analysis as a Mediator’s Tool*, 1 *Harvard Negotiation Law Review* 113–137 (1996); Howell E. Jackson, Louis Kaplow, Steven M. Shavell, W. Kip Viscusi, David Cope, *Analytical Methods for Lawyers*, 2003, at 1–33; Jeffrey M. Senger, *Decision Analysis: Decision Analysis in Negotiation*, 87 *Marquette Law Review* 723–735 (2004); Marc B. Victor, *The Proper Use of Decision Analysis to Assist Litigation Strategy*, 40 *The Business Lawyer* 617–929 (1985).

<sup>41</sup> Trees similar in general structure are given in Larry Laudan & Harry D. Saunders, *Re-Thinking the Criminal Standard of Proof: Seeking Consensus About the Utilities of Trial Outcomes*, 7 *International Commentary on Evidence* 1–34 (2009), at 5, and Reid Hastie & Robyn M. Dawes, *Rational Choice in an Uncertain World*, 2001, at 35.

<sup>42</sup> More generally, note that there can be more than two branches, depending on the number of available decision options.

corresponding to the two competing versions presented by the parties. The probability associated with each of these two states of nature,  $\Pr(H_i)$ , is indicated below the respective branch. Following the branches further to the right leads to terminal points, representing decision consequences  $C_{ij}$ . For example, following the decision branch  $d_p$  (finding for the prosecution), and then the  $H_d$ -branch, leads to the consequence  $C_{pd}$  (a false conviction). Associated with every consequence  $C_{ij}$  is a loss  $L(C_{ij})$ , shown on the far right-hand side. Sometimes, end points are represented by triangles or diamond-shaped notes. For simplicity, we leave this graphical notation aside.

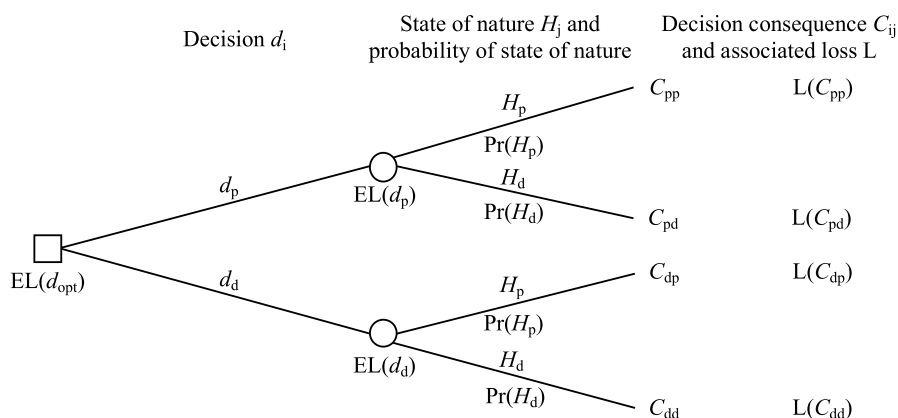
It is worth noting that the horizontal tree structure reflects several fundamental properties of decision-making under uncertainty, both in law and other areas of application:

- The decision-maker has control<sup>43</sup> *only* over the decision-branches that emanate from the trunk (squared node). She has *no* control, however, over the branches (representing states of nature) that grow out rightward from the (circled) chance nodes, and hence over decision consequences. It is unknown which state of nature represented by the latter branches actually holds, which is why the decision-maker can—at best—entertain probabilities for those states.
- The flow from the left to the right expresses a natural order in the sense that no decision consequence  $C_{ij}$  is incurred as long as no decision  $d_i$  is made.

So far, our decision problem has been described in its structure, but it has not yet been fully analysed. The decision tree merely displays all the possible ‘routes’ through which our actions and related consequences may unroll. The analysis, in turn, focuses on the computation of expected values, based on the assigned probabilities and losses. This is called ‘rolling back’ or backward analysis: starting from the terminal branches, expectations are computed at the chance points. For example, the expectation at the circled chance node of the upper branch in Fig. 1 is computed according to Equation (1):  $EL(d_p) = L(C_{pp})\Pr(H_p) + L(C_{pd})\Pr(H_d)$ . This value is the sum of the probability weighted losses of the two possible outcomes following decision  $d_p$  (i.e., finding for the prosecution). An analogous computation is performed for the chance point in the lower branch. Next, moving further back to the left, we reach the base (squared decision node). Below this node we write the expected loss of the optimal decision  $d_{opt}$ , which is understood here as the decision  $d_i$ ,  $i = \{p, d\}$ , which has the minimum expected loss.

This analysis conveys an insightful point: starting on the far right-hand side means to consider, first, the elements to happen *last* (i.e., decision consequences); in turn, the elements to occur first (i.e., decisions) are considered last. In the words of Lindley:

<sup>43</sup> “The only thing you can control is the decision (...)” (Howard (1983), *supra* note 34, at 93).



**Fig. 1** Decision tree for the current example in which  $d_p$  and  $d_d$  denote two available actions, finding for the prosecution and acquittal, respectively. These two decisions are represented as branches that emanate from the squared decision node shown on the left. Circled nodes represent chance points from which branches emanate to the right, representing uncertain states of nature (here: competing assertions  $H_p$  and  $H_d$  of the prosecution and defense, respectively). Decision consequences  $C_{ij}$ , resulting from deciding  $d_i$  when in fact  $H_j$  holds, and associated losses  $L(C_{ij})$  are shown on the far right hand-side of the tree, at its terminal points. Expected losses are indicated underneath the chance nodes. The expected loss of the optimal decision, i.e. the decision which has the minimum expected loss, is indicated below the decision node. Note that in a full analysis, the decision branch that does not have the minimum expected loss is double barred

The point is succinctly put. We cannot decide what to do today until we have decided what to do with the tomorrows that today's decisions might bring. This is an unfortunate, but indisputable, conclusion.<sup>44</sup>

This also seems a fair precept for legal applications as it is widely held that decision-makers ought to have a clear view about and appreciation of the possible decision consequences, especially related value judgments, *prior* to deciding.

### 2.3.2 Influence Diagrams

While decision trees lay out, from left to right, a natural time flow of the decisions and subsequent contingencies, influence diagrams are more compact representations that concentrate on reflecting the assumed relationships (dependencies) among the constituting elements of the decision problem. However, decision trees and influence diagrams are not opposing concepts, but complementary.<sup>45</sup> For example, a decision tree may be used to show the evaluation of an influence diagram, whereas an influence diagram may be used to synthesise a decision analysis where the decision tree becomes too 'bushy', a problem readily encountered when the number of possibilities that may result from decisions and chance events increases.

<sup>44</sup> Lindley (1971), *supra* note 39, at 148.

<sup>45</sup> French, Maule & Papamichail (2009), *supra* note 39, at 21.

Historically, literature on influence diagrams can be traced back to research conducted in the 1970s in the context of formal analyses of political conflicts and the use intelligence information gathered by governmental agencies.<sup>46</sup> Theoretical background of influence diagrams has mainly been published in the 1980s,<sup>47</sup> available also in more contemporary textbooks in both technical<sup>48</sup> and applied formats.<sup>49</sup>

As in decision trees, influence diagrams represent decisions and uncertain events (also sometimes called aleatory variables) by squares and circles, respectively. A third type of node, called value node, is diamond shaped and used to quantify the relative (un-)desirability of decision consequences using, for example, utilities or losses.<sup>50</sup> The nodes in a network are connected with arcs which can have different meanings. For example, an arc between two chance nodes in a network fragment  $A \rightarrow B$  indicates that the outcome of node  $A$  can influence the probabilities associated with  $B$ . In turn, arcs from decision and chance nodes to value nodes are called functional links. They express the fact that the possible consequences of a decision functionally depend on the possible states of nature and on the choice that is made among the possible decisions. The semantics of influence diagrams involve many more subtleties that go beyond the scope of this paper. We illustrate some further aspects directly in terms of an example.

The influence diagram corresponding to currently discussed example involving two decisions and two states of nature consists of three nodes connected as follows  $D \rightarrow L \leftarrow H$ . Compared to the decision tree shown in Fig. 1, one can see that there is only one node  $H$  (with two states  $H_p$  and  $H_d$ ) for the chance event, and one node  $L$  that contains the loss function  $L(d_i, H_j)$ , for  $d_i$  and  $H_j$  with  $i, j = \{p, d\}$ .

<sup>46</sup> Ronald A. Howard, James E. Matheson, Miley W. (Lee) Merkhofer, Allen C. Miller & D. Warner North, Comment on Influence Diagram Retrospective, 3 Decision Analysis 117–119 (2005).

<sup>47</sup> E.g., Ronald A. Howard & James E. Matheson, Influence Diagrams, in: The Principles and Applications of Decision Analysis, Vol. 2: Professional Collection, Ronald A. Howard & James E. Matheson (Eds.), 1983, at 719–762; Robert M. Oliver & James Q. Smith (Eds.), Influence Diagrams, Belief Nets and Decision Analysis, 1990 (Proceedings of the Conference entitled ‘Influence Diagrams for Decision Analysis, Influence and Prediction’, Engineering Systems Research Center, University of California at Berkeley, 1988); Ross D. Shachter, Evaluating Influence Diagrams, 34 Operations Research (1986), at 871–882.

<sup>48</sup> E.g., Robert G. Cowell, A. Philip Dawid, Steffen L. Lauritzen & David J. Spiegelhalter, Probabilistic Networks and Expert Systems: Exact Computational Methods for Bayesian Networks, 1999, at 155–188; Uffe B. Kjærulff & Anders B. Madsen, Bayesian networks and Influence Diagrams, A Guide to Construction and Analysis, 2008, at 74–91; Finn V. Jensen & Thomas D. Nielsen, Bayesian Networks and Decision Graphs, Second Edition, 2007, at 279–428; Richard E. Neapolitan, Learning Bayesian Networks, 2004, at 252–265.

<sup>49</sup> E.g., Robert T. Clemen & Terence Reilly, Making Hard Decisions with Decision Tools, 2001, 52–69; Monahan (2000), supra note 39, at 11–13; Kevin B. Korb & Ann E. Nicholson, Bayesian Artificial Intelligence, Second Edition, 2011; Kevin Murphy, Machine Learning: A Probabilistic Perspective, 2012, at 330–334; Stuart Russell & Peter Norvig, Artificial Intelligence: A Modern Approach, Third Edition, 2010, 626–636 (using the term “decision network”); Franco Taroni, Alex Biedermann, Silvia Bozza, Paolo Garbolino & Colin Aitken, Bayesian Networks for Probabilistic Inference and Decision Analysis in Forensic Science, 2014.

<sup>50</sup> Different scales may be used and, in some applications, it may be suitable to quantify costs, gains and rewards etc. directly in monetary terms.

This graph is considerably more compact, compared to decision trees, and represents the structural dependencies among the main constituting elements of the decision problem under study. It does *not* represent the temporal order of how decisions may be made and consequences may occur. Note also that the presence and direction of a particular arc is as informative as the absence of an arc. For example, there is no arc pointing from the decision node  $D$  to the chance node  $H$  because, in our example, the decision-maker's probability for the states of nature  $H_p$  and  $H_d$  are not affected by the decision made (i.e., whether deciding  $d_p$  or  $d_d$ ).<sup>51</sup>

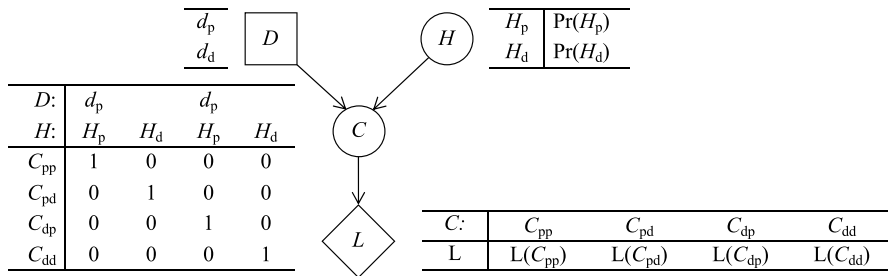
Although definitionally highly precise, influence diagrams are very flexible in the sense that the same problem can be represented by different structures and still lead to the same output. Structural variation may result, for example, from differences in the level of detail at which a decision problem is represented. Figure 2 shows an example of an alternative structure for the problem under study. Instead of conditioning the value node  $L$  directly on the decision node  $D$  and the chance node  $H$ , an intermediate node  $C$  is introduced. Its states represent all the consequences  $C_{ij}$  that result from decisions  $d_i$  and states of nature  $H_j$ . Associated with the node  $C$  is a node probability table. It contains the logical values 0 and 1. For example, the state  $C_{pp}$  is true and hence assigned the value 1, when  $d_p$  and  $H_p$  hold, and false (value 0) for all other combinations of  $d_i$  and  $H_j$ . In turn, the node  $L$  contains the loss  $L$  assigned to each possible consequence  $C_{ij}$ .

The influence diagram in Fig. 2 may be preferred to the diagram  $D \rightarrow L \leftarrow H$  when there is a need to express a clear distinction between consequences and the valuation of consequences. Regarding their output, however, the two models are equivalent: they both provide the expected loss for decisions  $d_p$  and  $d_d$  at the node  $D$ , in agreement with the formulaic result Equation (1), given the same probability distribution  $\Pr(H)$  and loss function  $L(\cdot)$ .

So far, our presentation of influence diagrams focused on static aspects, and the critical reader may ask how this graphical formalism may be useful in furthering the understanding of the decision theoretic concepts introduced so far. Whilst, at first sight, the elaborate graphical language may seem disproportionate with respect to the 'simple' output of the expected loss, corresponding to Equation (1), it is important to keep in mind that influence diagrams will be helpful for keeping track of the decision-theoretic treatment of more complex decision problems, for which purely formulaic developments or even decision trees will become too demanding. Translating non-formal arguments and lines of reasoning into a graphical representation using influence diagrams also helps to examine the logical structure of concepts, and their relationships, which are tasks that may be difficult to achieve without a formal framework. It is precisely this property of influence diagrams that we will exploit in later sections in order to examine the relationship between decision-theoretic and RP arguments.

<sup>51</sup> This may be different, for example, in a medical context where the probability of a patient's future health status may be affected by the decisions made regarding medical treatment.





**Fig. 2** Alternative influence diagram for a decision problem with two decisions  $d_p$  and  $d_d$ , and two states of nature  $H_p$  and  $H_d$ . For each combination of a decision  $d_i$  and a state of nature  $H_j$ ,  $i, j = \{p, d\}$ , a consequence  $C_{ij}$  is obtained. The states of the node  $C$  represent the possible consequences. The table associated with the value node  $L$  assigns a loss to each consequence  $C_{ij}$ , written  $L(C_{ij})$

Critics may also invoke that both decision trees and influence diagrams provide no solution to the fundamental problem of how to assign probabilities and values for decision consequences.<sup>52</sup> This is correct, of course, though it needs to be emphasized that the modelling frameworks do not claim to provide solutions for these challenges, nor do we intend to solve them in this paper. Rather, our premise is that decision-makers *are* uncertain about real-world events and that they *do* have preferences among decision consequences. Consequently, decision-makers are required to give an account of how they logically deal with these notions, which is a question in its own right, independently of how to quantify those notions in the first place. Specifically, as shown through Equation (2), numerical quantification is not a necessary requirement for decision theory to provide relevant insight: comparisons of (qualitative) orders of magnitude are sufficient.

### 3 Structural Relationships Between the Relative Plausibility and Decision-Theoretic Accounts for Criminal Adjudication

#### 3.1 Influence Diagram for Relative Plausibility Decisions

The question we now seek to investigate is (1) whether, and if so, how, the RP scheme<sup>53</sup> can be stated and analysed in terms of the graphical formalisms introduced in the previous section, and (2) which structural relationships between the RP and decision-theoretic accounts such analyses can reveal. This endeavour may seem incompatible with RP, because it is an explicitly non-formalistic concept. Notwithstanding, the understanding of an assessment or judgment of RP as a decision<sup>54</sup>

<sup>52</sup> Stated otherwise, in decision theory, it is assumed that the decision-maker can express probabilities and utilities (losses), for which the theory then provides instructions for their coherent combination.

<sup>53</sup> Throughout what follows, we mean RP as advocated by Professors Allen and Pardo (supra note 1).

<sup>54</sup> This assumption has been justified in §2.2.

should allow, at least in principle, a decision-theoretic analysis to be developed. Proponents of RP, we suggest, will agree that RP arguments should conform to principles of logical consistency.

To capture RP assessments using the language of decision theory, it is helpful to start with what the decision-maker is in control of: the RP *decisions*. As noted in §2, the decision-maker *judges* the plausibility of the parties' competing accounts. This means that, in essence, the decision-maker *must decide* which plausibility assessments most suitably address the parties' cases. This looks like there being two distinct decisions, one regarding the plausibility of the prosecution's account, and one regarding the defence's account. In essence, however, the plausibility judgments are—as their name suggests—relative, that is a given account's plausibility is not assessed in isolation, but against its competitor. Thus, RP can be regarded as an aggregate assessment. It represents a consideration that comes before the actual verdict. In this sense, there is, overall, a staged decision process, consisting of a (relative) plausibility assessment, followed by a verdict.

Based on the above, we identify the following three decisions available to the decision-maker in a typical criminal case:

$d_1$ : The prosecution (P) has no plausible account (PA).<sup>55</sup>

$d_2$ : P has a PA, and the defence (D) also has a PA.

$d_3$ : P has a PA, and D has no PA.

While decisions  $d_2$  and  $d_3$  clearly convey the compound nature of RP assessments, decision  $d_1$  does not explicitly mention the defence's account. That is, decision  $d_1$  reflects a situation in which the decision-maker considers that the prosecution was not able to present a sufficiently thorough account.<sup>56</sup> In practice, it is reasonable to consider, in accordance with the explanatory dimensions of RP, that the sufficiency of the prosecution case is being considered in relative comparison to the ordinary human experience of the trier of fact. This experience is the foundation for the interpretation of the criminal standard of proof. In the course of trial, influences on that experience derive from a positively presented defence account or defence challenges to prosecution evidence. In reviewing the evidence presented in conclusion of the trial, only if the prosecution has a plausible account, is there a practical impetus for the trier of fact to consider the nature of the defence account. It is important here not to place the prosecution and defence accounts in relative comparison because that would misconstrue the legal task. Both in law and practice, however, it may be observed that a plausible prosecution account is a precondition

<sup>55</sup> We acknowledge that in a 'typical' criminal case, the prosecution will have (almost by definition) a plausible account, at least from their own point of view, because otherwise there would not be a trial. This is, of course, not a conceptual problem, but a doctrinal and practical one. Jurisdictions have safety measures in place to prevent trials when the evidence against the defendant is weak or implausible, in order to raise efficiency (and save public money). Here, we include  $d_1$  as a decision in order to ensure generality of the argument.

<sup>56</sup> Unlike RP theory, conventional explanations of the criminal process would consider sufficiency as relative to the criminal standard of proof in law.

to any need for the trier of fact to separately consider the plausibility of an account proffered by a defendant.

In Professors Allen's and Pardo's RP theory, decision  $d_2$  leads to an acquittal, and only decision  $d_3$  leads to a guilty verdict.<sup>57</sup> A plausible account by the defence, in their scheme, can raise a reasonable doubt if it is compatible with innocence, not disproven or otherwise eliminated.<sup>58</sup> In summary, thus, there are two decisions,  $d_1$  and  $d_2$ , that lead to an acquittal, and only one decision,  $d_3$ , that is able to warrant a conviction.

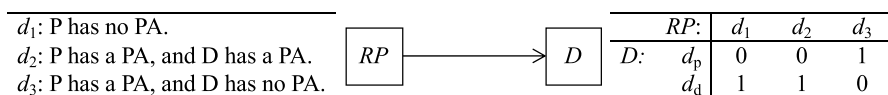
In terms of an influence diagram, the relationship between RP decisions and verdicts can be represented as shown in Fig. 3. The node *RP* has three states as defined above. They present all possible plausibility statements available to the decision-maker. Exactly one of the options needs to be chosen, and choosing one of them excludes the remaining options. The arc pointing from the node *RP* to the node *D* accounts for two aspects: a temporal order and a contingency. Chronologically, RP assessments precede verdicts. At the same time, they place a constraint on verdicts, which is expressed by the table associated with the node *D*. Following the constraints defined above, decisions  $d_1$  and  $d_2$  cannot warrant a verdict  $d_p$  in favour of the prosecution, but lead to an acquittal (decision  $d_d$ ): this is expressed by the values 0 (assigned to  $d_p$ ) and 1 (assigned to  $d_d$ ) in the first two columns. Column three contains the values 1 (assigned to  $d_p$ ) and 0 (assigned to  $d_d$ ). This expresses the view that a verdict in favour of the prosecution applies in a case in which the decision-maker considers that the defence does not have plausible account ( $d_3$ ). The node *RP* has no entering arcs, which reflects the idea that the decision-maker is not formally constrained in any way by other considerations. Though, in theory, the RP account acknowledges a wide range of considerations to inform plausibility assessments, these considerations are not formally structured at this stage of development of RP theory. Hence, they also have no formal representation in our influence diagram. In combination, these features reflect the highly liberal nature of RP theory.

The influence diagram in Fig. 4 is a minimalistic model in the sense that it only involves two decisions. It provides no guidance on how to choose a particular RP decision. It also incorporates no case specific characteristics. In essence, the RP model only specifies the possible decisions and their relationships.

This is different for standard decision-theoretic analyses where the aim is to look beyond decision points and investigate the parts of decision paths that are not under control of the decision-maker. This means to focus on uncertain events and uncertain decision consequences. Specifically, with respect to the influence diagram shown in Fig. 3, one may ask how to provide a criterion for choosing among the decisions  $d_i$ ,  $i = \{1, 2, 3\}$ , in view of the possible decision consequences, their probability and relative adversity. The modelling language of influence diagrams is ideally suited to investigate this question. We address it in the next section.

<sup>57</sup> Allen & Pardo, *supra* note 1, at 27.

<sup>58</sup> *Id.*



**Fig. 3** Influence diagram for the relationship between RP decisions, represented by the node *RP*, and verdicts (node *D*), i.e. decisions  $d_p$  and  $d_d$  in favor of the prosecution P and defence D, respectively. RP decisions are states of the node *RP* in the form of statements as to whether P and/or D has a plausible account (PA)

## 3.2 An Extended Influence Diagram for Relative Plausibility Decisions

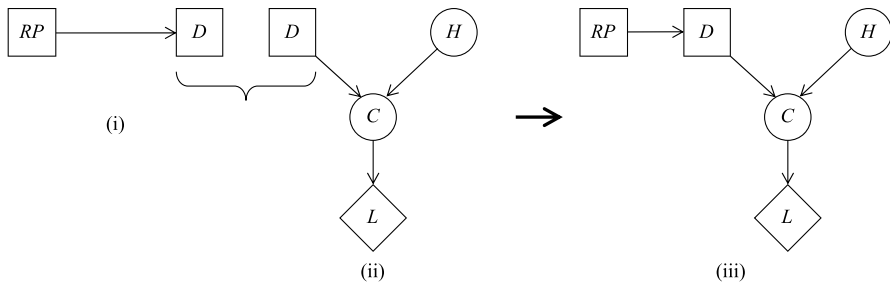
### 3.2.1 General Model Structure

So far, we have developed two influence diagrams in isolation, one for the classic decision-theoretic account according to Kaplan (§2.3.2), and one for the RP account (§3.1). At this point, these diagrams are disconnected, which raises the question whether there is a way to logically relate them to each other. This question is motivated by the understanding that the two concepts—optimality according to the decision-theoretic account and ‘best explanation’ in the RP account—appear to characterise distinct, but related aspects of the same overall process toward a decision regarding ultimate issues (i.e., verdict).

Inspection of the two influence diagrams suggests a natural way for combining them. In fact, both influence diagram fragments share *D* as a common node. A combined model can thus be proposed by retaining only one node *D*, as illustrated in Fig. 4. The resulting model 4(iii) exhibits a series of entirely reasonable qualitative properties. For example, consider that the loss is still related to only the node *C*, modelling the decision consequences. The node *L* is not directly impacted by the decision-maker’s plausibility assessment (node *RP*). The same holds for the decision consequences (node *C*): the scope of accurate and erroneous decision consequences is strictly a function of the verdict (node *D*) and the true state of nature (node *H*) only. This suggests that there should be no direct arc between the nodes *RP* and *C*. Notwithstanding, the decision-maker’s plausibility assessment is indirectly connected to the node modelling the consequences of the ultimate decisions: this connection is made through the path  $RP \rightarrow D \rightarrow C$ . In addition, because the table associated with the node *D* (see Fig. 4) defines a clear deterministic relationship between the decision nodes *RP* and *D*, it becomes possible to think about the decision consequences (node *C*) as a function of *RP*, and vice-versa: core features of the decision-theoretic sub-model, in particular probabilities for states of nature and losses for decision consequences, allow one to provide a measure to qualify and compare the various *RP*-decisions. The latter aspect is further explored below.

### 3.2.2 Analysis of Model Properties

To study the properties of the joint influence diagram (Fig. 4iii) in further detail, consider a situation in which the probabilities for the two competing accounts are  $\Pr(H_p)=0.3$  and  $\Pr(H_d)=0.7$ , respectively. Suppose further that an erroneous



**Fig. 4** Combination of the influence diagrams for RP, (i), and for the decision-theoretic account, (ii), through the common node *D*. The combined influence diagram, (iii), contains only a single node *D*. Definitions of all nodes and node tables are as given in Figs. 2 and 3

finding for the prosecution is the overall worst consequence, and thus is assigned the loss  $L(C_{pd})=1$ . An erroneous finding for the defence is considered ten times less adverse, and hence is assigned the loss  $L(C_{dp})=0.1$ . The most desirable consequences are correct findings for the prosecution ( $C_{pp}$ ) and an accurate acquittal ( $C_{dd}$ ): they are assigned zero loss, expressing the idea that nothing is ‘lost’ by taking an accurate decision.<sup>59</sup>

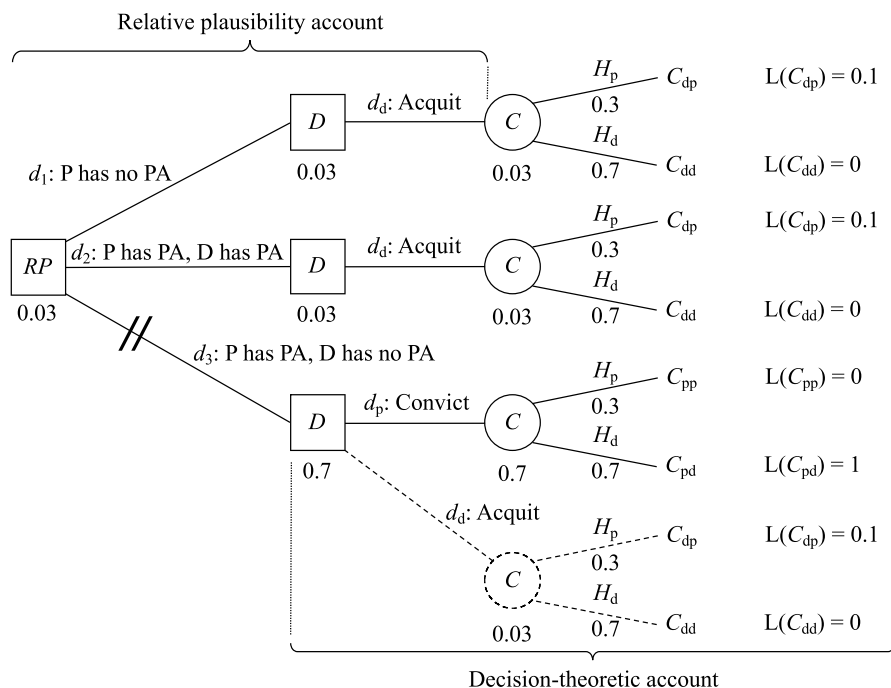
In the classic decision-theoretic account, the optimal decision is  $d_d$  (Acquit) because from Equation (1) it is readily seen that the expected loss for this decision,  $EL(d_d)$ , is 0.03, which is smaller than the expected loss for the decision  $d_p$  (Convict),  $EL(d_p)=0.7$ .<sup>60</sup> This is intuitively reasonable, given that the defence’s account is taken to be considerably more probable than the prosecution’s account. We can represent this decision analysis as shown in Fig. 5, within the wider context of a decision-theoretic account of RP. As mentioned previously, RP assessments precede ultimate decisions and thus extend the decision problem to what is called a two-stage decision problem. As an example, suppose we focus on the decision  $d_1$  (P has no PA),<sup>61</sup> a decision that looks well in line with the low probability of the prosecution’s account. Since the RP decision scheme stipulates an acquittal in a case in which the prosecution has no plausible account ( $d_1$ ), the expected loss  $EL(d_d: \text{Acquit})=0.03$  of decision  $d_d$  (Acquit) is carried over to the  $d_1$  decision branch.<sup>62</sup> Stated otherwise, we have obtained a decision-theoretic qualification of RP decisions. This is an important result precisely because it provides a criterion for comparing rival RP statements (i.e., decisions).

<sup>59</sup> The reader may choose other values, but should keep in mind that the general idea is to suppose that the decision-maker holds stronger beliefs in  $H_d$  than in  $H_p$  being true, so that  $\Pr(H_d) > \Pr(H_p)$ , and prefers accurate over erroneous decision consequences, thus assigning loss values accordingly.

<sup>60</sup> In fact, according to Eq. 1,  $EL(d_p) = L(d_{pp})\Pr(H_p) + L(d_{dp})\Pr(H_d) = 0 \cdot 0.3 + 1 \cdot 0.7 = 0.7$ , and  $EL(d_d) = L(d_{dp})\Pr(H_p) + L(d_{dd})\Pr(H_d) = 0.1 \cdot 0.3 + 0 \cdot 0.7 = 0.03$ .

<sup>61</sup> The analysis for other RP decisions is developed later in this section.

<sup>62</sup> As further explained below, the  $d_p$ -branch is not sketched because it is not an admissible decision under  $d_1$ .



**Fig. 5** Decision tree for the three RP decisions  $d_1$  (P has no PA),  $d_2$  (P has PA, D has PA) and  $d_3$  (P has PA, D has no PA), followed by decisions  $d_d$  (Acquit) and  $d_p$  (Convict), respectively. Definitions of all nodes and notation are as given in Figs. 2 and 3. Probabilities are noted below branches, and expected values below nodes. The branch of the non-optimal decision (i.e., the decision that does not minimize expected loss) is double-barred

In isolation, however, the expected loss  $EL(d_1: \text{P has no PA})$  is little informative. The crucial question is how it compares to the expected losses of the alternative decisions  $d_2$  (P has PA, and D has PA) and  $d_3$  (P has PA, and D has no PA). We can readily investigate these questions by inspecting the decision tree shown in Fig. 5, i.e. the branches for  $d_2$  and  $d_3$ , and the associated decision-analytic results.

Note that the branch corresponding to the decision  $d_2$  (P has PA, and D has PA) has the same structure as the branch for the decision  $d_1$  (P has no PA). The reason for this is that, as for the decision  $d_1$  (P has no PA), RP theory calls for an acquittal in a case in which decision  $d_2$  (P has PA, and D has PA) is made. The expected loss of an acquittal has already been determined as 0.03, which is thus carried over to the decision branch  $d_2$ . Although decisions  $d_1$  and  $d_2$  express fundamentally different views about the RP of the parties' cases, their expected losses are the same: this is because both RP decisions ( $d_1$  and  $d_2$ ) lead to an acquittal, which determines the applicable expected loss.

The third branch, corresponding to the decision  $d_3$  (P has PA, and D has no PA), is interesting in the context of this example because, in general, RP theory stipulates that when the prosecution has a plausible account, but not the defence,

then this is a case of conviction ( $d_p$ ). Could the decision-theoretic characterization of decision  $d_3$  show that this is not an optimal decision *in this case*? We ask this because, after all, the prosecution's case in our example is considered to have a *lower* probability than that of the defence. Under these circumstances, the expected loss for a conviction (decision  $d_p$ ) is:

$$\begin{aligned} EL(d_p) &= L(d_{pp}) \Pr(H_p) + L(d_{pd}) \Pr(H_d) \\ &= 0 \cdot 0.3 + 1 \cdot 0.7 = 0.7. \end{aligned}$$

This expected loss also characterizes decision  $d_3$  (P has PA, and D has no PA) because RP theory stipulates that  $d_3$  forms the basis for a conviction ( $d_p$ ). However, this expected loss, 0.7, is considerably higher than the expected loss of the rival RP decisions  $d_1$  and  $d_2$ . Thus, the decision  $d_3$  is non-optimal in decision-theoretic terms. We express this by double-barring the  $d_3$  decision branch. Since, in RP theory, an acquittal is not an option stipulated under decision  $d_3$  (P has PA, and D has no PA), the decision tree does not include it. However, we added it tentatively, using dashed lines, but merely to illustrate its expected loss. This allows us to show why the rival decisions  $d_1$  and  $d_2$  are better in this case: they are better because they stipulate an acquittal (decision  $d_q$ ) which, under the current circumstance (i.e., assigned probabilities and losses), has a lower expected loss. Further properties of the decision-analytic account of RP are presented in the Appendix.

## 4 Conclusions

Formal models of the legal process involving mathematical structures face a stream of critiques. Proposals of explicitly non-formalistic approaches, such as RP, emphasise greater versatility, but attract their own critiques, such as a lack of unequivocal and precisely articulable decision criteria. Some challenges apply to both mathematical and non-formalistic perspectives, such as the extent to which they can realistically be expected to find consistent application across different decision-making bodies.

While these methodological and conceptual divides seem insurmountable, they pose a fundamental challenge to both perspectives. On the one hand, RP cannot (afford to) dispense with logically defensible reasoning patterns altogether in favour of a vague intuition about what it means to say that a story is plausible. RP does not explicitly provide commonly understandable, i.e. linguistic criteria for its basic concept (plausibility). On the other hand, the mathematical nature of more formal perspectives remains a hurdle in practice given the holistic and qualitative tendencies of human reasoning processes. The feasibility of providing a logical architecture for RP assessments demonstrates how different perspectives can beneficially interact. We have sought to provide such an architecture by conceptualising and explaining RP assessments as decisions, using the conceptual framework of decision theory as a vehicle for capturing, analysing and conveying the logical relationships between the ostensible competing concepts of RP and classic decision theory. This structural analysis supports the view that adherence to one or the other of the examined

perspectives does not imply a contradiction, but reflects the coverage of different aspects of the same overall decision process. Our architecture also preserves the RP property of not being bound by any sort of external measure against which particular accounts are to be assessed. In our analysis, RP assessments represent observable surface features of decision-making behaviour that explain the grounds of verdicts (conviction and acquittal). Notwithstanding, these ultimate issues inevitably remain—in their nature—decisions under uncertainty that can be formally characterised in classic decision theoretic terms.

There is no prescriptive suggestion, at this point, for the abstract and theoretical decision structures we have described to be quantitatively implemented in current legal proceedings. They are merely means to gain insight into the nature of the problems that evidence and proof disciplines confront, and for discussants in those disciplines to rationalize what they are doing.

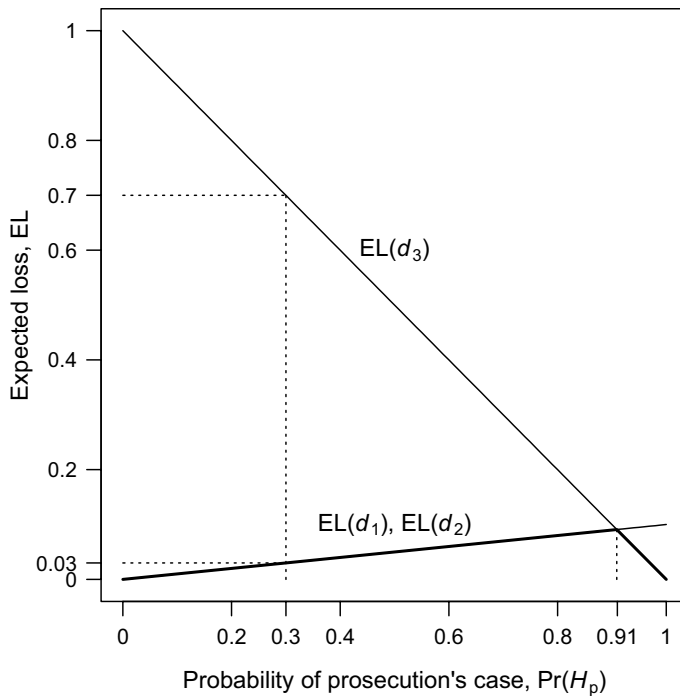
## Appendix: Further Properties of the Decision-Theoretic Account of Relative Plausibility

Consider again the example presented in §3.2.2. Fig. 6 shows the expected losses for the RP decisions  $d_1$ ,  $d_2$  and  $d_3$  not only for the current case where the probabilities for the prosecution's and defence's accounts are, respectively, 0.3 and 0.7, but for the full range of values between 0 and 1. As may be seen, RP decisions  $d_1$  (P has no PA) and  $d_2$  (P has PA, and D has a PA) are optimal, i.e. have the minimum expected loss, for a broad range of probabilities, including probabilities greater than 0.5 (i.e., when the probability of the prosecution's account is greater than that of the defence).

There is a transition point, however, toward the right-hand side of Fig. 6. For probabilities greater than this change point, the expected loss of decision  $d_3$  (P has PA, D has no PA) is smaller than that of the alternative decisions  $d_1$  and decision  $d_2$ . This threshold probability corresponds to the minimum probability necessary in the classic decision theoretic account to ensure that the decision  $d_p$  (finding for the prosecution; conviction) has a *smaller* expected loss than the decision  $d_d$  (Acquit), given a particular loss ratio as specified by Equation (2). Recall that in the case here, an erroneous finding for the prosecution ( $C_{pd}$ ) is considered ten times worse than an erroneous acquittal ( $C_{dp}$ ). Thus, following Equation (2), the odds defining the transition point are 10:1, corresponding to  $\Pr(H_p) = 0.91$ .

It is important to keep this result in mind because when looking at Fig. 6, the skeptical reader may ask how it can be possible that for a probability of the prosecution's account,  $\Pr(H_p)$ , as high as 0.8, or even 0.9, the decisions  $d_1$  (P has no PA) and  $d_2$  (P has PA, and D has a PA) have a *lower* expected loss, and hence are preferable to decision  $d_3$  (P has PA, and D has no PA). As noted above, the explanation for this observation stems from the chosen loss function, in particular the ratio of the loss associated with an erroneous finding for the prosecution ( $C_{pd}$ ) and the loss associated with an erroneous acquittal ( $C_{dp}$ ). In the case here we have chosen, for the sole purpose of illustration, a ratio of 10:1. Hence, a finding for the prosecution is not warranted, in decision theoretic terms, for





**Fig. 6** Expected losses (on y-axis) of RP decisions  $d_1$  (PA has no PA),  $d_2$  (PA has PA, and D has PA) and  $d_3$  (PA has PA, and D has no PA) as a function of the probability of the prosecution's case (x-axis),  $\Pr(H_p)$ , for the example discussed in the text. The dotted vertical line at  $\Pr(H_p)=0.3$  indicates the expected losses for decisions  $d_1$  and  $d_2$ , i.e. 0.03, and decision  $d_3$ , i.e. 0.7. Note that these values correspond to the values illustrated in the decision tree shown in Fig. 5. The dotted vertical line at  $\Pr(H_p)=0.91$  indicates the transition point where, for probabilities greater than this value, decision  $d_3$  (PA has PA, and D has no PA) has a smaller expected loss, and hence is better than the alternative decisions  $d_1$  and  $d_2$ . The bold line indicates, for every probability  $\Pr(H_p)$  between 0 and 1, the decision(s) with the minimum expected loss

situations in which the probability for the prosecution's case,  $\Pr(H_p)$ , is smaller than 0.91. Case examples with loss ratios so that values of  $\Pr(H_p)$  as high as 0.8 or 0.9 would be sufficient for the RP decision  $d_3$  (P has PA, and D has no PA) to be optimal in a decision-theoretic sense can be found in Table 1. More generally, note that we do not suggest, at this point, that a full numerical quantification be imposed on practical RP decisions. The sole point we seek to make is that it is possible to give a formal (mathematical) justification for the *intuition* that the higher the stakes involved (i.e., the more one of the two ways of deciding erroneously is considered worse than the other), the lower should be one's quantum of doubt.

As a last example, consider a case in which the prosecution's account  $H_p$  is considerably more probable than the account presented by the defence,  $H_d$ . Specifically, let  $\Pr(H_p)$  be 0.95. Table 2 summarises a few examples of cases with different loss ratios (columns 1 to 3). The optimal RP decision and related verdict for each

**Table 2** Extension of Table 1 with optimal RP decision(s) (in column 6) and associated verdict (in column 7) for a hypothetical case in which the decision-maker's probability for the prosecution's account,  $H_p$ , is 0.95

$L(C_{pd})$	$L(C_{dp})$	Loss ratio ( $x$ )	Limiting value $\Pr(H_p)$	Current belief $\Pr(H_p)$	Optimal RP decision	Verdict
1	1	1	0.5	0.95	$d_3$	$d_p$
1	0.5	2	0.66	0.95	$d_3$	$d_p$
1	0.1	10	0.91	0.95	$d_3$	$d_p$
1	0.01	100	0.99	0.95	$d_1$ or $d_2$	$d_d$
1	0.001	1000	0.999	0.95	$d_1$ or $d_2$	$d_d$

The possible RP decisions are:  $d_1$  (PA has no PA),  $d_2$  (PA has PA, and D has PA) and  $d_3$  (PA has PA, and D has no PA). Decisions  $d_p$  and  $d_d$  represent the verdicts 'convict' and 'acquit', respectively. Column 4 provides the minimum probability necessary, in the classic decision-theoretic account, to ensure that the expected loss for deciding for the prosecution,  $d_p$ , is smaller than the expected loss for deciding in favour of the defence,  $d_d$ , given different loss ratios (column 3). Note that a loss ratio  $x > 1$  means that the loss incurred by an erroneous decision  $d_p$  (column 1) is  $x$  times greater than the loss incurred by an erroneous decision  $d_d$  (column 2). We assume that an erroneous finding for the prosecution, i.e. a wrongful conviction,  $C_{pd}$ , is the overall worst consequence and thus assigned the loss 1

example is given in columns 6 and 7. Note, in particular, that for loss ratios such as 100:1, or more, a current belief of 0.95 is *not* sufficient to warrant the RP decision  $d_3$  (P has a PA, and D has no PA).

As an aside, note that throughout this paper, all examples considered a loss ratio  $x \geq 1$ , which means that the loss incurred by an erroneous decision  $d_p$  (i.e., wrongful conviction) is at least equal or greater than the loss incurred by an erroneous decision  $d_d$  (i.e., erroneous acquittal). Consideration of opposite cases seems unnecessary given current understandings of the values upheld in contemporary legal orders.

We also emphasise the important conclusion that there is no single and absolute probability to warrant a particular RP decision  $d_3$  (P has a PA, and D has no PA) and hence a finding for the prosecution (decision  $d_p$ , conviction). In our decision-theoretic account of RP, the minimum probability necessary to warrant a conclusion that the prosecution has a plausible account (and that the defence does not have a plausible account) is found by weighing the odds of the two competing accounts against the losses that may be incurred by the two ways in which a verdict may turn out erroneously. This offers an answer to the recurrent critique of the RP theory that it does not provide an explicit criterion for decision-makers to determine *when exactly* the prosecution's account has reached the required level of proof and, in addition, the defences' account is insufficient to raise a reasonable doubt.

**Publisher's Note** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

## Affiliations

**Alex Biedermann<sup>1</sup>**  · **David Caruso<sup>2</sup>** · **Kyriakos N. Kotsoglou<sup>3</sup>**

✉ Alex Biedermann  
alex.biedermann@unil.ch

<sup>1</sup> Faculty of Law, Criminal Justice and Public Administration, School of Criminal Justice, University of Lausanne, 1015 Lausanne-Dorigny, Switzerland

<sup>2</sup> University of Adelaide Law School, The University of Adelaide, Ligertwood Building, North Terrace, Adelaide 5005, Australia

<sup>3</sup> Northumbria Law School, Northumbria University, Newcastle upon Tyne NE1 8ST, UK