DOI: 10.1111/risa.14037

PERSPECTIVE

A risk science perspective on liability/guilt and uncertainty judgements in courts

Terje Aven | **Roger Flage**

University of Stavanger, Stavanger, Norway

Correspondence
Terje Aven, University of Stavanger, Stavanger, Norway.
Email: terje.aven@uis.no

Abstract

This article aims to provide new insights about risk and uncertainty in law contexts, by incorporating ideas and principles of contemporary risk science. The main focus is on one particular aspect of the law: its operation in courts where a defendant has been charged with a violation of civil or criminal law. Judgements about risk and uncertainty—typically using the probability concept—and how these relate to the evidence play a central role in such situations. The decision on whether the defendant is liable/guilty or not may strongly depend on how these concepts are understood and communicated. Considerable work has been conducted to provide theoretical and practical foundations for the risk and uncertainty characterizations in these contexts. Yet, it can be argued that a proper foundation for linking the evidence and the uncertainty (probability) judgements is lacking, the result being poor communication in courts about risk and uncertainties. The present article seeks to clarify what the problems are and provide guidance on how to rectify them.

KEYWORDS
knowledge strength, probability concept, risk and law

1 | INTRODUCTION

Consider a court where a defendant has been charged with a violation of civil or criminal law. In a civil case, the defendant is found liable or not liable for damages, whereas in a criminal case the decision is guilty or not. For simplicity, in this article, we only use the terms guilty or not guilty. What are the consequences of delivering a verdict of guilty or not guilty? Basically, there are four: (1) the defendant did not commit the violation that he/she is accused of but is found guilty, (2) the defendant committed the violation but is not found guilty, (3) the defendant did not commit the violation and is not found guilty, and (4) the defendant committed the violation and is found guilty. The two first outcomes represent errors and are commonly referred to in statistics as errors of type I and type II, respectively.

It is, however, not known with certainty whether the defendant actually committed the violation. There are uncertainties and, hence, risk.

The above paragraphs express the essence of the concept of risk, as defined by current risk science as understood by the authors of the present paper (e.g., SRA, 2015). It has two main elements: (i) the consequences of the activity considered with respect to something that is of value to us and (ii) the associated uncertainties. The activity here is the trial with its verdict.

However, when people and lawyers discuss risk in such contexts, is this the way this concept is understood? No, this is not a typical perspective. More common is that they associate risk with probabilities: the probability of the defendant being found guilty when not having committed the violation, as well as the probability that the defendant is found not guilty when having committed the violation, in other words, the probability of errors of types I and II. In statistics, probability has precise interpretations, but, in law contexts, as here discussed, the conceptual clarity is often weak, and other more specific ‘‘law interpretations’’ are also seen, based on links to the evidence and judgements of its strength, through the concept of ''burden of proof.'' This concept is the obligation to present evidence on the subject of the lawsuit or the criminal charge. In criminal cases, the requirement is that all accusations against the defendant be proven ''beyond a reasonable doubt.'' The standard is that there is a large likelihood that the accused committed the crime. The burden is on the prosecution because of the initial presumption of innocence. Probability numbers are not specified, but levels like 90%, 95%, or even 99% are often referred to (Newman, 2019; Shoener, 2014; see also Section 2.2).

In civil cases, the standard instruction of the proof burden states that an affirmative defense must be established by a preponderance of the evidence, where ''preponderance'' means more likely than not (Allen & Stein, 2013). The formulation leads directly to a probabilistic interpretation of larger than a 0.5 probability. The standard also refers to a claim of ''balance of probabilities,'' with the same interpretation.

Although these standards relate to the concept of likelihood and probability, it is not trivial to explain what probability or likelihood means in this context. The issue has been thoroughly discussed in the literature (see, e.g., Allen & Stein, 2013; James, 1941; Pardo & Allen, 2008). On the one hand, we see perspectives which seek to integrate the burden of proof reasoning into statistical and decision analysis frameworks (e.g., Bell, 1987; Lindley, 1977; Orloff & Stedinger, 1983), like the Bayesian one; on the other, we find perspectives which make a clear distinction between such frameworks and the law context. Allen and Stein (2013) provide an example of the latter type of perspectives. These authors refer to the concept of relative plausibility, which is understood as "instructing factfinders to determine which of the parties' conflicting stories makes most sense in terms of coherence, consilience, causality, and evidential coverage" (Allen & Stein, 2013). They argue that the concept corresponds to our courts' practice is in line with people's natural reasoning and common sense, and outperforms mathematical probability, operationally and normatively. Allen and Stein (2013) assert that our courts apply mathematical probability only to a small number of well-defined categories of cases. While not an area of extensive research in the risk science literature, there are examples of both broader discussions, such as on the relationship between the fields of law and risk science (Harlow, 2004), and on analysis of more specific risk-related concepts in a law context, such as risk and statistical evidence (Brannigan et al., 1992).

The likelihood and probability concepts are essential for the risk and uncertainty communication in courts and, ultimately, for the decision to be made. However, reading the literature on this topic reveals that the foundation for the risk and uncertainty communication in this context is weak. Current discussions only to a limited degree acknowledge and reflect that probability and likelihood judgements cannot be considered isolated from the knowledge and evidence supporting these judgements. In this article, we argue that contemporary risk and uncertainty science provides insights which can be used to provide increased clarity on the nexus between the concepts of risk, uncertainties, knowledge, evidence, and probability (likelihood) in this law context. First, in Section 2, we provide two cases to illustrate the discussion. Then, in Section 3, we present a framework that clarifies the meaning of and relationship between these concepts. The framework is discussed in Section 4, and, finally, Section 5 provides some conclusions.

2 | TWO CASES TO ILLUSTRATE THE DISCUSSION

2.1 | A civil case

We consider a case where a person, the Plaintiff, was involved in a rear-end collision. He sued for damages. The accident was relatively minor, but the Plaintiff claimed that it caused some severe back problems and pain. The defendant, on the other hand, argued that the crash was not the cause of the Plaintiff's pain issues; rather, these could be traced back to some previous back condition of the Plaintiff. A case like this is described by Magraken (2010). The Plaintiff's allegation of causation of the ongoing injuries due to the collision was dismissed by the Court, which stated

> It may well be that Mr. X's continuing back pain is the result of soft tissue injuries resulting from the accident. However, on the balance of probabilities, Mr. X has failed to persuade me that this is the case… (Magraken, 2010)

Thus, the court refers to probability and gives this justification, by writing

> Dr. Y's carefully worded report really says it all: "[t]he soft tissue injuries were responsible for his symptoms and limitation soon after the accident" [emphasis added] and, "[t]he injuries sustained were not severe enough to aggravate his pre-existing conditions." I note that it is possible that Dr. Y is wrong. It could be that the whiplash from the accident did affect Mr. X's spine in a way that affected the area of the surgery. But without a full and proper medical investigation with the aid of diagnostic imaging, I have no way of knowing that. The body of evidence that has been put before me on behalf of Mr. X, who bears the burden of proving his case on the balance of probabilities, just does not do that (Magraken, 2010).

Let A denote the event or statement that the accident ''caused'' the pain problems, and P the probability of the event occurring (the statement being true), i.e., $P = P(A)$. The conclusion in the case was that the court could not conclude that $P > 0.5$, given the evidence and knowledge available, and hence it has to be dismissed, according to the "balance of probabilities" standard. Let K denote this evidence and knowledge. Clearly, there is a link between K and P, but, as discussed in Section 1, it is not straightforward and has been subject to considerable discussion in the literature. In Section 3, we will present a framework which aims at contributing to clarifications. The key is to be precise on the meanings of P and K.

2.2 | Criminal law: ''Beyond reasonable doubt''

Judge Weinstein is a leading authority on the American law of evidence (Tillers & Gottfried, 2006). He advocates for probabilistic reasoning in relation to criminal law and interpretation of "beyond reasonable doubt," using statements like "a probability of guilt of no less than 95% should be necessary to support a conviction" and "in the high 90s" (Weinstein & Dewsbury, 2006). Other law scholars have formulated similar suggestions, see for example Franklin (2006), Tillers and Gottfried (2006), and Newman (2019).

Weinstein and Dewsbury write:

> In a criminal case, the law tilts in favour of defendants; it prefers that some guilty go free rather than that some innocents be convicted. The questions are (1) how high the required minimum probability should be set and (2) how should the test be articulated. Quantification is one way of stating the standard. Requiring a combined descriptive explanation and an explicit percentage would likely work well for the thousands of jurors we have observed (Weinstein & Dewsbury, 2006, p. 167).

Weinstein and Dewsbury emphasize that the precise meaning of the standard, that is, how high it is, is left to the jury, reflecting different views among the jurors on for example what are acceptable risks of wrongful convictions and acquittals. Influencing factors in this regard may include (Weinstein & Dewsbury, 2006):

- Jurors may alter the probability required for conviction based on their degree of confidence in the police, the prosecutor, the court, and the justice system as a whole. Feelings about such matters can vary significantly from one community to another.
- Jurors may require a lower standard of proof based on their perception of the danger of the moment.
- The jury may elect to convict a defendant for the crime charged on a lower standard of proof based on evidence of defendant's past crimes or problematic life style.
- Jurors may be influenced by their beliefs about the future dangerousness of the defendant.

According to Weinstein and Dewsbury (2006), an approach based on such probabilistic statements would have "the advantage of drawing the jurors' attention to the high standard of proof in criminal cases, as well as the law's policy of avoiding conviction of the innocent." In addition, the approach points "the jurors towards the need to consider absence of proof that would be expected were the investigation thorough" (Weinstein & Dewsbury, 2006).

3 | OUTLINE OF A CONCEPTUAL FRAMEWORK

Before presenting the framework, we provide some reflections related to the above cases and use of probabilities.

A key question is how to interpret the statement $P > 0.5$, that the probability exceeds 50% (or, e.g., $P \geq 0.95$). The literature commonly refers to probability theory and different perspectives on how to understand a probability: the frequentist perspective and the subjective perspective. The former reflects variation and is defined as the fraction of times the event under consideration occurs, if the situation considered could be repeated over and over again infinitely. It is a mind-constructed quantity which has to be estimated. The latter perspective expresses the uncertainty or degree of belief of the assessor that the specific event will occur, or the statement is true. In the setting discussed here, the issue is whether a statement is true or not—and the subjective probability is the only relevant perspective. However, in the law literature, it is seldom that this is stated explicitly, with a clear interpretation of what a subjective probability means. This is surprising, as it is essential for understanding a statement like $P > 0.5$, and its relationship to the evidence and knowledge K. To establish a proper basis for characterizing uncertainties and risk, we need to base it on well-defined concepts with interpretations. There are two main challenges in relation to such characterizations.

The first relates to how to best transform the evidence and knowledge available to probability. The key issue is to what degree it is possible to obtain an "objective transformation," not influenced by the assessor's subjective judgements. The risk and uncertainty sciences have thoroughly discussed this issue; the main conclusion is that the use of precise (crisp) probability acknowledges that the probability assignment is in fact subjective and includes assessor judgements, as reflected by the term "subjective probability," and that imprecise probabilities reduce the subjectivity and allow for more objective characterizations (Flage et al., 2014); see below for more details. The use of imprecise probability is more objective but less informative. In the extreme case, the adoption of imprecise probability may lead to an interval of the type [0,1]; thus, the assessor is not willing to say anything about the probability of the statement being true or the event occurring.

The second challenge relates to the strength of the evidence and knowledge. The probability specified can be based on a weak evidence or knowledge basis. How should this be taken into account in the uncertainty judgements?

3.1 | Basic concepts

The definitions and explanations presented are based on SRA (2015).

3.1.1 | Probability

Let A be the statement or event of interest, and $P(A)$ the probability of A. This probability is "subjective" or "knowledge-based" and interpreted with reference to a comparison standard, as explained by these examples:

$P(A) = 0.50$ means that the assessor has the same uncertainty or degree of belief in the statement A being true or the event A occurring, as randomly drawing a red ball out of an urn containing 100 balls, of which 50 are red.

$P(A) > 0.50$ means that the assessor has the same uncertainty or degree of belief in the statement A being true or the event A occurring, as randomly drawing a red ball out of an urn containing 100 balls, of which 51 or more are red—the assessor is not willing to be more precise. The probability is referred to as an imprecise probability.

From these examples, it is straightforward to provide interpretations for any precise or imprecise probability. We omit the details.

The probability $P(A)$ is conditional on knowledge K; for short, we write $P(A) = P(A|K)$.

3.1.2 | Uncertainty

For the law context considered in this article, uncertainty relates to whether or not the defendant committed the violation that he/she is charged with. In the example of Section 2.1, this amounts to whether the accident "caused" the pain problems or not. The situation is unique or, in other words, not repeatable. At the decision point, the court does not know what the truth is. It has a lack of knowledge. It is faced with uncertainties in the sense of lack of knowledge. For the assessment of uncertainties, see Second 3.1.4.

3.1.3 | Knowledge and evidence

Knowledge is of different types, but for the present discussion it is best explained as "justified beliefs." For the car example of Section 2.1, the court has knowledge about the issue being discussed, a main source being the doctor's report. Knowledge is not the "truth" but beliefs—claims—about the truth. The justification thus becomes critical, and scientific practice is much about how to perform the justification, for example, using statistical methods and treating uncertainties. In general, the knowledge is founded on data, information, modeling, testing, argumentation, so on. Often, the knowledge adopted in an assessment can be formulated as assumptions. When referring to the knowledge K, we commonly also include the data, information, modeling, testing, argumentation, so on that build this knowledge. Knowledge is thus used in both a narrow sense, as justified beliefs, and a broad sense, as justified beliefs and the just-mentioned elements.

It is essential to distinguish between general knowledge (GK) and specific knowledge (SK) for the activity considered. To illustrate, consider the example of Section 2.1. GK includes all the generic medical knowledge available on the issues discussed, whereas SK includes knowledge concerning the specific situation, related to the accident and Mr. X, with all his history and characteristics.

In addition, we need to clarify whose knowledge is referred to. We distinguish here between the court's knowledge, Dr. Y's knowledge, and the total knowledge available when also adding other experts' knowledge.

Evidence as a concept is closely related to knowledge. Basically, it captures the basis for a belief or statement (here, A), in the form of data, information, modelling insights, test results, analysis results, so on. Schematically, we can write:

Knowledge (broad sense) = Knowledge (narrow sense)
+ Data, information, modelling, testing, argumentation, etc.
= Justified beliefs + Evidence.

The probability $P(A|K)$ can then be rewritten as $P(A|B,E)$, where A is the event of interest, B is a set of justified beliefs, and E is the evidence. The event of interest, principally violation or not violation in the context considered here, is then measured by a degree of belief. The justified beliefs, B, on the other hand, are taken for granted, that is, are considered to be true.

A risk assessment can produce evidence in the form of a risk characterization of relevance for making a judgement about the statement A being true. The doctor's report in the example can be viewed as a risk assessment, and the court uses it as evidence, acknowledging that it is not representing the truth but a judgement about the truth subject to uncertainties (see Table 1).

3.1.4 | Risk

The basic ideas of the risk concept are introduced in Section 1. Risk is related to the future consequences of an activity and has basically two features: the consequences, with respect to the values of interest (violation or not, lives, environment, material assets, etc.) and uncertainties (what will the consequences be?). When considering risk, there is always a potential for negative or undesirable consequences, but, at the time considered, we do not know what the result of the activity will be; it could also be positive.

To describe or characterize how large the risk is, we enter the risk assessment sphere, which includes a specification of the consequences (e.g., by focusing on the events, violation or not) and an assessment of the uncertainties. The assessment of the uncertainties leads to a characterization of the uncertainties of the form (Q,K), where Q is a measure or description of the uncertainties, and K, the knowledge that Q

TABLE 1 Illustrations of the fact that knowledge at a specific level (Dr. Y) can be evidence at the next level (court)

Actors	Statement or event	Knowledge (justified beliefs)	Evidence
The court	The accident ''caused'' the pain problems		Dr. Y's statement: "[t]he injuries sustained were not severe enough to aggravate his pre-existing conditions"
Dr. Y		[t]he injuries sustained were not severe enough to aggravate his pre-existing conditions	Medical check-up

is based on. For Q, it is recommended to use probability (precise or imprecise), together with a judgement of the strength of the knowledge K.

3.2 | Dealing with the two main challenges of uncertainty and risk characterizations

In this section, we will present risk science knowledge on how to deal with the two challenges introduced at the beginning of Section 3.

3.2.1 | Challenge 1: The transformation from knowledge to probabilities

For law applications, imprecise probabilities are commonly used and are justified by the fact that such probabilities are sufficiently precise for its purpose, as what is needed is $P > p$ or $P \geq p$, where p is 0.50 or a number close to 1 (e.g., 0.95); refer to the discussion in Sections 1 and 2. However, these probability statements are still to some degree dependent on the assessor. If we write $P(A|K)$ to highlight the knowledge and evidence supporting the assignment P, there is a subjective leap from K to P. Consequently, it would not make sense to rely on the probability statements alone: K and its strength always need to be part of the risk and uncertainty characterizations and discussion.

This recognition of imprecise probabilities does not, however, mean that precise probabilities cannot be used as important evidence. For example, a probability judgement of say 0.90 can be made by some experts expressing their view on a matter, based on their knowledge. The number is acknowledged for what it is: a subjective judgement made by these experts, conditional on their knowledge. Assuming that they are highly qualified experts, the decision makers could give considerable weight to this judgement.

3.2.2 | Challenge 2: The strength of the knowledge supporting the probabilities

The assessor specifies an imprecise probability, for example $P > 0.50$ or $P \geq 0.95$, and adds a judgement of the strength of this knowledge. It is important to know whether the knowledge supporting the assignment $P > 0.50$ or $P \geq 0.95$ is strong or weak. But how should this judgement be made? We recommend using a qualitative categorization, for example, based on three categories: strong, medium strong, and weak knowledge. In the criminal law example of Section 2.2, we may think of two situations, (a) and (b).

In situation (a), there is a lot of scientific evidence as well as eyewitness evidence supporting the proposition that the defendant committed the crime rather than that somebody else committed the crime. In (b), there is only one eyewitness testimony and no scientific evidence supporting the proposition that the defendant committed the crime rather than that somebody else committed the crime. Suppose that in both cases a probability of minimum 0.95 is assigned. The cases differ however on the strength of the supporting knowledge and evidence. It is clearly stronger in the former case than in the latter, and this is expressed by classifying for example the knowledge as strong in the case (a) and weak in case (b).

The literature commonly refers to the likelihood ratio $P(\text{evidence } |V)/P(\text{evidence } |\text{not } V)$, where V is the event that the accused did commit the violation, when discussing evidence or knowledge strength (see, e.g., Lindley, 1977). Cases with stronger evidence and knowledge would typically have a large ratio compared to cases with weak knowledge. However, this way of measuring evidence and knowledge strength has some limitations. First, it is not clear what the probability of the evidence means. To make sense, the evidence must be an event or a statement which can be true or false, but evidence cannot in general be given such a form. But even if it can be given such a form, this would not in general solve the problem. For cases (a) and (b), we can think about situations where all the evidence is in the form of statements which are considered to be true. Hence, the ratios become 1 in both cases. Still, even if the probabilities can be properly interpreted, there is also the issue of the knowledge supporting these probability assignments. How should we reflect differences in this knowledge?

We recommend using a categorization of the strength of the knowledge supporting P relying on an intuitive understanding of these terms (strong, medium strong, weak knowledge/evidence), but, to improve consistency, some guidance for the specification is needed. Some criteria should preferably be introduced. First, the data and information supporting the knowledge K need to be looked into. How relevant are the data and information available for the issue considered? How much relevant data and information do we have? Another criterion relates to the degree to which we understand the phenomena considered. In the example of Section 2.1, the judgement of the Court was that the understanding was rather weak, without performing a proper medical investigation. As a third criterion, we need to evaluate key assumptions made. To what degree are they reasonable? What would be the effect of making changes in one or more of these assumptions? As a fourth criterion, the degree of agreement or consensus among experts is considered, in particular for experts coming from different "schools." As a final and fifth criterion, we recommend addressing the degree to which the knowledge has been thoroughly examined, for example, in relation to potential surprises. Aspects to consider include (Bjerga & Aven, 2016):

- The possibility of unknown knowns (i.e., others, but not the analysis group, have the knowledge). Have special measures been implemented to check for this type of event (e.g., the use of an independent review of the analysis)?
- The possibility that events are disregarded because of very low probabilities, although these probabilities are based on critical assumptions. Have special measures been implemented to check for this type of event (e.g., signals and warnings influencing the existing knowledge basis)?

Based on judgements of such issues, qualitative score systems can be developed; see, for example, Flage et al. (2014), Aven and Flage (2018), and Aven (2017). We also refer to the so-called NUSAP system (NUSAP: Numeral, Unit, Spread, Assessment, and Pedigree), which is based on similar ideas to reveal the knowledge strength (Berner & Flage, 2016; Funtowicz & Ravetz, 1990, 1993; Kloprogge et al., 2011; Laes et al., 2011; van der Sluijs et al., 2005a, 2005b). See Mosleh and Bier (1996) and Flage et al. (2014) for conceptual clarifications of the related issue of uncertainty about probability.

If knowledge is decomposed into justified beliefs and evidence, a more nuanced strength of knowledge scheme can be developed, by developing criteria related to:

(1) The strength of evidence: How good is the evidence that is involved?
(2) The strength of justification: How well supported are the justified beliefs by the evidence?

Returning to the car case of Section 2.1, the available evidence (the medical investigation) may be judged as strong if the investigation has followed proper medical procedures; however, the strength of justification for the statement ''[t]he injuries sustained were not severe enough to aggravate his pre-existing conditions'' may not be judged as strong, due to the lack of diagnostic imaging forming part of the evidence.

4 | DISCUSSION

In the car example of Section 2.1, the court can be seen as concluding that $P > p$ for p less than 0.5, but it was not willing to state that $P > 0.5$, as the evidence and knowledge supporting $P > 0.5$ was too weak. We see that the court's reasoning can be well placed in the framework of Section 3. The framework presented allows for referring to both probabilities and judgements of evidence and knowledge.

It is common to use the Bayesian approach for integrating the evidence and the uncertainties (Lindley, 1977). To describe this approach, consider the conditional subjective probability $P(A|K_0)$, where K_0 is the evidence as presented to the court, that is, the probability that the accident ''caused'' the pain problems, given the evidence K_0. The Bayesian approach is based on the use of Bayes' formula, stating that

$$P(A|K_0) = P(K_0|A)\,P(A)/P(K_0).$$

Here, $P(A)$ refers to the prior probability of the event A, that is, a judgement about the statement "The accident 'caused' the pain problems", before considering the evidence. A possible choice is to set this probability equal to ½, to aim at neutrality. Another possible choice is to set the prior probability equal to the fraction f of cases in which similar accidents cause similar pain problems, assuming that such a fraction is known (to the medical community). Now, suppose that we knew that the accident "caused" the pain problems, what is then the probability that the evidence turned out as it did? This is the probability that $P(K_0|A)$ seeks to reflect. However, the meaning of this probability is not clear. As discussed in Section 3.2, to make sense, the evidence K_0 must be an event or a statement which can be true or false, but evidence cannot in general be given such a form. In this case, the key evidence was a report from Dr. Y, providing a statement and rationale for concluding that the injuries sustained were not severe enough to aggravate Mr. X's pre-existing conditions. If B denotes the event that Dr. Y concludes that the injuries sustained were not severe enough to aggravate his pre-existing conditions, it is possible to give an interpretation of $P(K_0|A) = P(B|A)$: the probability that the doctor concludes that the accident was not the cause, when in fact it was the cause. Similarly, we need to assign the probability of $P(B|$ not $A)$, that is, the probability that the doctor correctly concludes that the accident was not the cause, when that is in fact the case. It is, however, difficult to specify the probabilities $P(B|A)$ and $P(B|$ not A), with some justification. The issue is about how good the doctor is in his profession, in respect of this particular issue. Clearly, $P(B|A)$ is a small number, whereas $P(B|$ not $A)$ is close to one. The probabilities will depend on the general knowledge on the matter. In this particular case, GK can be considered rather strong, but there are still some uncertainties and discussions about the effects of this type of event (whiplash).

As seen above in relation to Bayes' formula, the probability $P(A)$ can be linked to the relative frequency f of similar events to A. If such a frequency can be introduced, in the sense that similar situations exist, but this frequency is unknown, the probability P can be assigned through the following reasoning: If the value of the frequency f were known, we would assign this value as the probability of

A, that is, we would assign $P(A|f) = f$. This is what was presented as an option in relation to Bayes' formula above. On the other hand, if the value of the frequency f is unknown, we can use subjective probability to describe the uncertainty about f. Specifically, we can assign a probability distribution to f. Then, it follows from the rules of probability that $P(A) = E[P(A|f)]$, where E denotes an expected value, evaluated with respect to the probability distribution on f. Such an approach relies on the idea of similar events. Technically, the requirement is that the situations are exchangeable, meaning roughly that if a given number of similar situations are considered, it is only the number of occurrences of the event among these situations that decide the probability and not the specific situations in which these events occur. Exchangeability, and thus whether two or more situations are similar, is a judgement by the person assigning the probability. So, although the procedure for assigning the probability by introducing similar situations is objective, the premise and starting point for using this procedure involves a subjective judgement. In addition, the (precise) probability distribution expressing uncertainty about f is a subjective probability judgement.

Relating the concept of strength of knowledge to probability naturally raises the question of whether stronger knowledge as a rule gives a higher probability P. This may indeed be the case, as will be seen in the example in the next paragraph. However, it does not have to be the case, as illustrated by the following example: We return to the Bayes' formula example above. Having Dr. Y's statement as part of the knowledge base of a probability judgement regarding the event A (that the accident ''caused'' the pain problems) makes for a stronger knowledge base than not having this statement. Now, consider as suggested above a prior probability $P(A) = ½$. Given B (the doctor concludes that the accident was not the cause), either $P(A|\mathrm{B})$ or $P(\text{Not } A|B)$ will decrease as the sum is 1. The stronger knowledge intuitively gives a higher probability $P(\text{not } A|B)$; hence, $P(A|B)$ will decrease. Use of Bayes' formula will show that this holds as long as $P(B|A)/P(B|\text{not } A) <1$. Suppose for example that $P(B|A) = 0.05$ and $P(B|\text{ not } A) = 0.99$. Then, it follows from Bayes' formula that the updated (posterior) probability of A, given B, is

$$P(A|B) = P(B|A)\,P(A)/P(B) = P(B|A)\,P(A)$$
$$/\left[P(B|A)\,P(A) + P(B|notA)\,P(notA)\right]$$
$$= 0.05 \times 1/2\left[0.05 \times 1/2 + 0.99 \times 1/2\right] = 0.048.$$

The probability that the accident "caused" the pain problems has thus been substantially reduced, by approximately a factor 10, due to the increased and stronger knowledge provided by Dr. Y's statement. Of course, if Dr. Y were judged to be less reliable, which would be reflected by a greater value of $P(B|A)$ and/or a smaller value of $P(B|\text{ not } A)$, this conclusion could change. For example, it is seen that if both $P(B|A) = 0.5$ and $P(B|\text{ not } A) = 0.5$, then the statement by Dr. Y has no effect on the updated probability. In this case, both the numerator and each of the two denominator terms are equal to 0.5 × 0.5, and we get $P(A) = P(A|B) = ½$.

Suppose that the Plaintiff appeals against the court's ruling and that the case is moved to a higher court. Also suppose that, as part of the case preparation, the Plaintiff consults another medical expert, say Dr. Z, who requests a diagnostic imaging study and, based on its findings, concludes that the area of the past surgery has damage consistent with an external impact and with the type of pain reported by the Plaintiff. Let K_1 denote the evidence that is presented to the higher court, where K_1 then includes the statement from Dr. Z. In this case, the higher court may conclude that $P(A|K_1) > p$ for p greater than 0.5, as a result of increased knowledge.

5 | CONCLUSIONS

This article brings ideas and principles from contemporary risk science into the law context, specifically into the problem of a court reaching a verdict on the guilt or not of a defendant. While the risk and probability terms are frequently used and there is a considerable amount of literature on the use of probability in this context, we conclude that the understanding and the use of these concepts are not in line with what the authors of the current article consider as contemporary risk science. A main point highlighted is that probabilities (precise and imprecise), which express uncertainties and degrees of belief, cannot be seen isolated from the knowledge and evidence supporting the probability, and in particular the strength of this knowledge and evidence. As a contribution to improve the situation, we describe a conceptual framework, placing law on a modern risk science foundation. The framework covers the key concepts, including probability, uncertainty, knowledge, evidence, and risk, as well as recommendations on how to deal with some main challenges of uncertainty and risk characterizations. Compared to previous frameworks in other contexts, a main emphasis is placed on clarifying the relationship between evidence and knowledge. While current risk science acknowledges that uncertainty and risk characterizations are always based on some knowledge, in a law setting it is particularly useful to distinguish between knowledge in a narrow sense, understood as justified beliefs, and knowledge in a broad sense, also including evidence, understood as the basis of these beliefs. This leads to three different elements involved when assessing uncertainty using probability:

first, the event of interest, principally the question of violation or not of a law, assessed using subjective probability (precise or imprecise) expressing a degree of belief;

second, the justified beliefs, conceptualized as statements that are taken for granted, that is, considered as true, that the degree of belief assignments are based on;

and, finally, the evidence that the justified beliefs are based on, which could include data, information, modeling insights, test results, analysis results, so on.

The distinction between knowledge in a narrow and a broad sense has implications for strength of knowledge assessments that need to supplement uncertainty assessments. Such assessments have previously been made directly on the overall body of knowledge. With the distinction here made, we can distinguish between strength of justification for knowledge in a narrow sense and strength of evidence. Strength of knowledge assessment then refers to the combined assessment of these two aspects.

For interpreting the "balance of probabilities" and "beyond reasonable doubt," we have argued that it is not sufficient to refer to probabilistic statement of the form $P > 0.50$ and $P \geq 0.95$, we also have to relate these judgments to the strength of the supporting knowledge and evidence. In the "beyond reasonable doubt" case, also the supporting knowledge needs to be strong. This framework is based on the legal principle of treating defendants as innocent until proven guilty which requires not only a large probability P, but also strong supporting evidence and knowledge. For the case of "balance of probabilities," we conclude that the supporting knowledge for $P > 0.50$ should be stronger than that supporting $P \leq 0.50$, reflecting that an affirmative defense must be established by a preponderance of the evidence.

ACKNOWLEDGMENT

The authors are grateful to three reviewers for their useful comments and suggestions to earlier versions of this article.

ORCID

Roger Flage https://orcid.org/0000-0002-2967-9104

REFERENCES

Allen, R. J., & Stein, A. (2013). Evidence, probability, and burden of proof. *Arizona Law Review*, *55*, 557–602.

Aven, T. (2017). Improving risk characterisations in practical situations by highlighting knowledge aspects, with applications to risk matrices. *Reliability Engineering & System Safety*, *167*, 42–48.

Aven, T., & Flage, R. (2018). Risk assessment with broad uncertainty and knowledge characterisation. In T. Aven & E. Zio (Eds.), *Knowledge in risk assessment and management* (pp. 1–26). John Wiley & Sons.

Bell, R. S. (1987). Decision theory and due process: A critique of the Supreme Court's lawmaking for burdens of proof. *Journal of Criminal Law & Criminology*, *78*, 557–585.

Berner, C. L., & Flage, R. (2016). Comparing and integrating the NUSAP notational scheme with an uncertainty based risk perspective. *Reliability Engineering & System Safety*, *156*, 185–194.

Bjerga, T., & Aven, T. (2016). Some perspectives on risk management: A security case study from the oil and gas industry. *Journal of Risk and Reliability*, *230*(5), 512–520.

Brannigan, V. M., Bier, V. M., & Berg, C. (1992). Risk, statistical inference, and the Law of Evidence: The use of epidemiological data in toxic tort cases. *Risk Analysis*, *12*(3), 343–351.

Flage, R., Aven, T., Zio, E., & Baraldi, P. (2014). Concerns, challenges, and directions of development for the issue of representing uncertainty in risk assessment. *Risk Analysis*, *34*(7), 1196–1207.

Franklin, J. (2006). Case Comment—United States v. Copeland, 369 F. Supp. 2d 365 (E.D.N.Y. 2005): Quantification of the 'proof beyond a reasonable doubt' standard, 5 L. *Probability & Risk*, *159*, 159–165.

Funtowicz, S. O., & Ravetz, J. R. (1990). *Uncertainty and quality in science for policy*. Kluwer Academic Publishers.

Funtowicz, S. O., & Ravetz, J. R. (1993). Science for the post-normal age. *Futures*, *25*, 735–755.

Harlow, S. D. (2004). Science-based trade disputes: A new challenge in harmonizing the evidentiary systems of law and science. *Risk Analysis*, *24*(2), 443–447.

James, G. F. (1941). Relevancy, probability and the law. *California Law Review*, 29, 689–705.

Kloprogge, P., Van Der Sluijs, J. P., & Petersen, A. C. (2011). A method for the analysis of assumptions in model-based environmental assessments. *Environmental Modelling & Software*, *26*(3), 289–301.

Laes, E., Meskens, G., & van der Sluijs, J. P. (2011). On the contribution of external cost calculations to energy system governance: The case of a potential large-scale nuclear accident. *Energy Policy*, *39*(9), 5664–5673.

Lindley, D. V. (1977). Probability and the law. *Journal of the Royal Statistical Society: Series D (The Statistician)*, *26*(3), 203–212.

Magraken, E. (2010). Proving Your BC Injury Claim: A 'Balance of Probabilities'—BC Injury Law. https://bc-injury-law.com/proving-bc-injury-claim-balance-probabilities/

Mosleh, A., & Bier, V. (1996). Uncertainty about probability: A reconciliation with the subjectivist viewpoint. *Systems, Man and Cybernetics, Part A: Systems and Humans*, *26*(3), 303–310.

Newman, J. O. (2019). Taking "beyond a reasonable doubt" seriously. *Judicature*, *103*(2), 33–41. https://judicature.duke.edu/articles/taking-beyond-a-reasonable-doubt-seriously/

Orloff, N., & Stedinger, J. (1983). A framework for evaluating the preponderance-of-the-evidence standard. *University of Pennsylvania Law Review*, *131*(5), 1159–1174.

Pardo, M. S., & Allen, R. J. (2008). Juridical proof and the best explanation. *Law and Philosophy*, *27*(3), 223–268.

Shoener, N. (2014). Burden of proof in civil and criminal cases. Legal Match Law Library. https://www.legalmatch.com/law-library/article/burden-of-proof-in-civil-and-criminal-cases.html

SRA. (2015). Society for Risk Analysis glossary. https://www.sra.org/resources/

Tillers, P., & Gottfried, J. (2006). A collateral attack on the legal maxim that proof beyond a reasonable doubt is unquantifiable. *Law, Probability and Risk*, *5*, 135–157.

van der Sluijs, J., Craye, M., Futowicz, S., Kloprogge, P., Ravetz, J., & Risbey, J. (2005a). Combining quantitative and qualitative measures of uncertainty in model-based environmental assessment. *Risk Analysis*, *25*(2), 481–492.

van der Sluijs, J., Craye, M., Funtowicz, S., Kloprogge, P., Ravetz, J., & Risbey, J. (2005b). Experiences with the NUSAP system for multidimensional uncertainty assessment in model based foresight studies. *Water Science and Technology*, *52*(6), 133–144.

Weinstein, J. B., & Dewsbury, I. (2006). Comment on the meaning of 'proof beyond a reasonable doubt'. *Law, Probability and Risk*, *5*, 167–173.

How to cite this article: Aven, T., & Flage, R. (2022). A risk science perspective on liability/guilt and uncertainty judgements in courts. *Risk Analysis*, 1–8. https://doi.org/10.1111/risa.14037