ORIGINAL PAPER

Reasonable Doubt, Robust Evidential Probability and the Unknown

Hylke Jellema[1]

Abstract

Most legal evidence scholars agree that proof of guilt beyond a reasonable doubt requires the belief that the defendant probably committed the alleged acts. However, they also agree that this is not a sufficient condition, as this belief may be unreasonable. I focus on two popular proposals for additional conditions: (i) that the degree of belief should be *robust* and (ii) that it should be reasonable given the available evidence (should be an *evidential probability*). Both criteria face difficulties regarding their meaning and utility. I propose an interpretation which overcomes these difficulties and which combines the two conditions into a single criterion on which both conditions are about the possibility of overlooked, exculpatory information. Proof of guilt beyond a reasonable doubt requires that we are sufficiently certain that we did not miss such information for the sake of achieving *accurate beliefs* and a *just error distribution*. How certain we can be that we did not miss anything relevant depends on our *higher-order evidence*. One important type of higher-order evidence is whether the hypotheses under consideration are sufficiently *detailed*. This point leads to a novel way of combining the *Bayesian* and *explanation-based* accounts of rational legal proof.

Keywords Legal epistemology · Criminal proof · Bayesianism · Explanation-based accounts · Robustness · Probability

✉ Hylke Jellema
   h.jellema@rug.nl

1   Faculty of Law, University of Groningen, Oude Kijk in 't Jatstraat 26, Groningen 9712 EK, The Netherlands

Springer

1 Introduction

When is guilt proven beyond a reasonable doubt? There is wide-spread agreement among evidence scholars that one requirement is that the defendant (very) probably committed the alleged acts. There are various ways in which we could interpret the word 'probable'. The most common interpretation is the Bayesian account on which probabilities are interpreted as the fact-finder's (judge or jury's) degree of belief (Hunt & Mostyn, 2020). For example, suppose that proving murder requires the defendant to have killed someone with intent and premeditation then the judge or jury should not convict him unless they believe that he (very) likely did so. However, a high degree of belief cannot be all there is to proof of guilt beyond a reasonable doubt. For instance, we may believe something strongly but may recognize that our evidence does not support this belief. Second, the available evidence may support a strong belief, but this set of evidence may be defective. For instance, many crucial items of evidence may be missing or the process of collecting the evidence may have been biased. In both cases, the fact-finder is arguably not justified in concluding that guilt has been proven.

In response to such worries, evidence scholars have put forward proposals for further requirements for proof of guilt beyond a reasonable doubt. In this article, I focus on two popular ones. The first is that the degree of belief should be robust (or resilient, safe or stable) (Ho, 2008, 278; Stein, 2005, 88; Di Bello, 2013; Dahlman et al., 2015; Urbaniak, 2018). Briefly put, on this requirement the fact-finder's belief should not be easily overturned. The second is that the fact-finder's high degree of belief should be reasonable in the face of the available evidence (that it should be an 'evidentially calibrated', 'reasonable', 'epistemic', or 'evidential' probability) (Nance, 2016; Wittlin, 2019; Spottswood, 2019; Hedden & Colyvan, 2019).

Both criteria face difficulties. For the most common interpretation of robustness, the problem is that it is unclear why it is worthwhile to have robust beliefs. For evidential probability the worry is that the notion is so vague that it does not tell us when one's belief is actually calibrated to the evidence. In Sect. 2, I offer a new interpretation of both terms which overcomes these difficulties. On my account both criteria are about the same issue: the need to take into account the possibility that we have overlooked exculpatory information.

Of course, we can never be certain that we did not overlook anything, nor can we know whether what we may have overlooked was exculpatory. Nonetheless, we can be justified in *presuming* that we considered all pertinent information. If this presumption is not justified there is reasonable doubt, or so I shall argue (Sect. 3). More precisely, too great a possibility of missed exculpatory information should lead to a reasonable doubt for the sake of two central goals of criminal trials: *error minimization* and *error distribution*. Error minimization means that criminal fact-finding should lead to as few errors as possible. The goal of error distribution is that, to the extent that errors are made (as is unavoidable in a legal system), these errors should mostly be false acquittals, not false convictions.

Whether the presumption that we did not overlook anything is justified depends on our *higher-order evidence* – which I take to be evidence about the reliability of the fact-finder's inferences I discuss how such evidence relates to our first-order beliefs

and give some examples of higher-order evidence in Sect. 4. In Sect. 5 I suggest that an important type of higher-order evidence is how specific the hypotheses under consideration are. Thinking in terms of well-specified hypotheses helps ensure that the belief in the defendant's guilt is robust and supported by the available evidence. This observation leads to a new way in which the two most popular accounts of rational criminal proof, the Bayesian and explanation-based, can be combined (Sect. 5).[1]

2 Robust Beliefs that Respect One's Evidence

The evidential and robustness conditions are primarily amendments to the Bayesian account of rational legal proof. On this account, agents have 'degrees of belief' or 'credences' with respect to any proposition which express how strongly the agent believes in the truth of this proposition. Such degrees of can be modeled as a probability distribution, satisfying the standard axioms of the mathematical theory of probability. Additionally, the agent should conditionalize (update) their degrees of belief upon acquiring new information by way of Bayes' theorem. On the Bayesian account, the beyond a reasonable doubt standard is usually understood in terms of the fact-finder's degree of belief, where proof of guilt requires a high degree of belief that the defendant committed the alleged acts conditional on the evidence (e.g. at least 90%, 95% or even 99%) (Gardiner, 2019). However, some have worried that treating such a degree of belief as a sufficient condition would lead to an account that is 'too subjective', as it makes proof of guilt purely a matter of personal belief (Laudan, 2006, 80; Risinger, 2013, 71; Allen, 2017, 138). But beliefs can be unreasonable. For instance, we may believe something that the evidence does not support or the set of evidence on which we must base our belief can be insufficient quality to draw any conclusions from (e.g., because it is biased or too incomplete). The requirements of evidential probability and robustness target these distinct ways in which the fact-finder's degree of belief can be unjustified.

2.1 Evidential Probability

Recently, a number of evidence scholars suggest that to count as rational, an agent's degrees of belief should be 'calibrated to the evidence', 'reasonable', 'epistemic', or 'evidential' probabilities (Nance, 2016; Wittlin, 2019; Spottswood, 2019; Hedden & Colyvan, 2019). On this probability interpretation, the high probability required for proof of guilt beyond a reasonable doubt refers to the degree of belief that is *reason-*

---

[1] Let me briefly discuss a methodological point. Like many (if not most) authors who discuss the concepts of robustness and evidential probability, my account is not tied to any particular legal system. The aim of this article is to clarify the aforementioned concepts and to explain certain epistemic intuitions (e.g., the benefit of making one's hypotheses detailed). My arguments rely on the assumptions that criminal proof is about determining (probable) truth and that the aim of the criminal standard of proof is to distribute errors fairly. These assumptions are widely shared among legal evidence scholars. Nonetheless, the implications of the account proposed here *will* depend in part on the details of one's legal system. In particular, in some systems it may be up to the defense to raise the worry of undiscovered exonerating information before the fact-finder may consider this possibility. In other systems the fact-finder may play a more active role with respect to this possibility.

able to hold in the face of the available evidence. In other words, it is the degree of belief that a reasonable agent with the same evidence *would* have – regardless of whether the fact-finder actually holds this credence (Hacking, 2001, 130; Williamson, 2002, 209). However, most legal evidence scholars who discuss the notion of evidential probability do not offer an account of when this is the case. As Williamson (2002, 211), one of the most prominent authors who discusses evidential probability, points out, such an account may not be needed. According to Williamson (2002, 209), even if we cannot give a precise account of what evidential probability is, we may often perfectly intelligibly ask: "how probable is [a given hypothesis] on present evidence?" He therefore suggests that we should regard evidential probability as a primitive – a concept which cannot be spelled out in terms of more fundamental concepts, but which is clear enough in context.

While I agree that we may sometimes reasonably ask what conclusions the evidence objectively supports, there will also be cases in which the fact-finder *does* need more guidance than the slogan 'just look at the evidence'. For instance, suppose that two jurors or two judges disagree over whether the evidence supports the belief that the defendant is guilty. How might they settle this disagreement? Or imagine that a fact-finder believes that the defendant is probably guilty, but is not entirely certain about whether this belief conforms to the evidence. How should they then decide on whether guilt is proven beyond a reasonable doubt? In both cases we need a more precise account of how the fact-finder should determine whether their belief conforms to the evidence. However, existing epistemic frameworks may be too limited in scope to capture the evidential richness of legal proof, or so vague that they are uninformative (Redmayne, 2003). Yet without a clearer notion of how we should determine what our evidence supports, we risk reverting back to a strictly subjective interpretation of the proof standard, because we might end up at an account where evidential probability is just whatever conclusions the fact-finder believes are justified given the evidence (Acree, 2021, 273).

Is it possible to give a clearer account of evidential probability? In this article I propose a way of spelling out this idea in terms of whether the fact-finder has overlooked exculpatory information. This may seem like a strange suggestion at first: evidential probability relates to how probable a given hypothesis is in the light of information that we *do* have. So why would we want to analyze this concept in terms of unknown information? My answer to this question begins with the observation that an evidential probability is a probability assessment that is reasonable given the available evidence. However, in legal cases, what the relevant evidence is and what this evidence supports is not simply given. The relevant information needs to be gathered, selected and interpreted. Only once we have done so can we intelligibly ask what this evidence entails – i.e., what the evidential probability of our hypotheses is. I propose that we have assurances that our probability indeed aligns with the evidence to the extent that we have made sense of the evidence well. Admittedly, giving a precise and complete definition of what it means to make sense of the evidence well is a task that goes beyond my abilities and I am not certain that a fully satisfactory account is possible. Nonetheless, I want to suggest that it at least involves not overlooking relevant information.

Making sense of the evidence often boils down to 'seeing how the facts hang together'. A large part of making sense of the evidence well is therefore seeing relevant connections. These connections can be between the evidence and the hypothesis, but also between the items of evidence themselves. For instance, if witnesses agree with one another in their testimonies, this will typically have an impact on the combined strength of these testimonies. Furthermore, apart from seeing connections between the evidence, making sense of the evidence at the very least also involves coming up with hypotheses that explain the evidence. For instance, there may be alternative versions of what happened which are consistent with the defendant's innocence. To give an example of this, suppose that in a murder case there is, apart from the defendant, another person who could have realistically killed the victim. Our evaluations of the evidence can be unreasonable if we overlook such explanations. For instance we may then overvalue the strength of the case for guilt. Additionally, we can be unreasonable because we overlook inconsistencies within the explanations that we have come up with. To give an example, a particular scenario may seem plausible, but only because we have overlooked that it implies that the defendant traveled from point A to point B in an unrealistically short time – the hypothesis is then internally inconsistent. Finally, our evaluations of the evidence become more reasonable to the extent that we do not overlook relevant arguments. Such arguments include objections to the background beliefs that we use in the evaluation of the evidence and hypotheses. These beliefs can be unreasonable, for instance if they are contradicted by scientific evidence. I want to suggest that to make sense of the evidence adequately involves (at least) conceiving of relevant alternative hypotheses, noticing what evidence is relevant, seeing relevant connections between the items of evidence as well as any inconsistencies within the hypotheses under consideration and becoming aware of pertinent arguments for or against one's beliefs. When we overlook such information we have not adequately made sense of the evidence.

To summarize, I want to suggest that what the fact-finder needs to be assured of in order to reasonably believe that their belief conforms to the evidence is that they have not missed any alternative explanations for the evidence, connections between the items of evidence or arguments against their conclusions. If the fact-finder can reasonably believe that they did not overlook any exculpatory[2] information, then they can also reasonably believe that they have made sense of the evidence well. As I discuss in Sects. 4 and 5, we can meaningfully reason about this question by looking at our 'higher-order evidence'.

---

[2] In this article I am mainly concerned with the question whether exculpatory information was missed. But why should we not consider the possibility of having missed inculpatory information too? The reason for this is that I discuss the notion of evidential calibration as an additional condition for proof of guilt beyond a reasonable doubt, next to the requirement of the fact-finder having a high degree of belief in guilt. The question is when a lack of evidential calibration should lead to reasonable doubt. And for this, we consider whether information was missed that could have overturned the fact-finder's high degree of belief in guilt – i.e., exculpatory information.

2.2 Robustness

A second worry for the Bayesian interpretation of proof beyond a reasonable doubt is that, even if our belief lines up with the evidence, this set of evidence may itself be flawed. For example, this set may be too incomplete or it may have been collected in a biased way. If we *had* to draw a conclusion based on this evidence, we might, for instance, have to conclude that the defendant's guilt is highly probable. However, many share the intuition that such a conclusion unwarranted if based on a flawed set of evidence. For this reason, various authors propose that one's belief should be *resilient* or *robust*, which refers to the stability of one's degrees of belief. However, there are actually two ways to understand this idea, which are sometimes mixed up. The first is robustness as sensitivity to new information (i.e., 'how would our belief change supposing that we were to find further information supporting innocence?'). The second is robustness as the absence of undiscovered exonerating information (i.e., 'is it realistic to presume that a more thorough search could have discovered information that would have overturned our belief in guilt?'). The first of these interpretations is problematic and should be rejected. The second is more plausible and *should* be adopted as a condition for proof of guilt beyond a reasonable doubt.

2.2.1 Robustness as Insensitivity to New Information

The notion of robustness as insensitivity to new information was introduced to explain how evidential *weight* is reflected in one's degrees of belief. Weight refers to how substantial the set of evidence is on which we base a specific belief (Keynes, 1921, 77). For example, suppose that you find a coin. You do not know whether the coin is biased or fair. Imagine that you are considering the probability that the one-hundredth flip of this coin will land on heads. Because you know nothing about the coin, your credence in this hypothesis is 0.5. Now you toss the coin ninety times and find that it does indeed land on heads roughly as often as it lands on tails. As a result, you come to believe that the coin is fair. Your credence that the 100th flip will land on heads therefore remains 0.5. Nonetheless, the two situations are different. In the second case, the credence of 0.5 is based on a greater amount of information. Hence, while your credence has stayed the same, the weight on which that credence is based has increased.

Skyrms (1977) and others suggest that weight is reflected in the robustness of our beliefs, where robustness is a measure of how much our credence changes in the face of additional data. For example, consider the coin flipping situation again. Suppose that you flip the coin nine times and all nine times it lands on heads. If these were the first nine flips, your belief that the coin is biased towards heads should increase significantly and therefore so should your belief that the hundredth toss will be heads. However, if you have already flipped the coin ninety times, you have a substantial amount of data which suggests that the coin is not biased. Hence, flips ninety-one to ninety-nine all landing on heads should not influence your belief in the outcome of the one-hundredth flip as strongly as in the situation where you did not flip the coin before these nine flips. So, your belief is more robust than in the first case and this is the result of the second probability being based on a weightier set of evidence.

Some suggest that robust beliefs should be a requirement for proof of guilt beyond a reasonable doubt (Logue, 1997; Ho, 2008; Stein, 2005; Dahlman et al., 2015; Dahlman & Nordgaard, 2023; Urbaniak, 2018; Di Bello, 2013). However, while it may have some intuitive plausibility, the robustness condition is problematic, as it is unclear why having robust beliefs is epistemically valuable and, hence, why a lack thereof should lead to reasonable doubt. Nance (2016, 270-8) surveys the literature on robustness and finds no convincing arguments in favor of this position. Additionally, Hamer (2012) argues that while the robustness condition was introduced to explain how evidential weight is reflected in one's beliefs, in legal proof, having a more weighty set of evidence may actually correlate with *less* robust beliefs. As we saw in the coin-flip example above, if we perform a series of independent, identical trials then increasing the weight of the evidence will also increase its robustness. In other words, weight and robustness covary – as one goes up or down, so does the other (and vice versa). However, the same is not necessarily true in criminal proof. Criminal investigations are not about conducting repeated, identical trials. Instead, investigators gather a diverse set of evidence. Yet as the weight, and therefore diversity of our evidence increases, the robustness of our beliefs may decrease (Hamer, 2012). This happens because more investigation often means that we make our hypotheses more specific. We then develop more detailed, fine-grained accounts of what might have happened. Such specific hypotheses are less resilient because it is easier to overturn a specific statement than a general one. For example, consider the following hypotheses:

> The defendant was at the crime scene between 9 a.m. and 9:15 a.m.
> The defendant was at the crime scene between 9 a.m. and 5 p.m.

Now suppose that we gather evidence which indicates that the defendant was not at the crime scene before 10 a.m. This falsifies the first hypothesis, but not the second. In the first case we might therefore have to conclude that the evidence does not support the belief that the defendant is (probably) guilty. In the second case the evidence does not lead to this conclusion. So, at least sometimes, a reasonable fact-finder obtaining more evidence will lead to their beliefs becoming less robust. This is worrying for the defender of a robustness condition, as optimizing weight is arguably an epistemically valuable thing to do. For instance, to support this claim Nance (2016) refers to a theorem by Horwich (1982, 127-9) in the philosophy of science which shows that as we accrue relevant evidence, our expected error – which is the difference between our credence in a proposition and its truth value – goes down in the long run. Others, such as Allen & Pardo (2007, 134) and Stein (2005, 122-3) note informally that it generally seems to be the case in legal proof that as the amount of evidence increases, factual error decreases. But if robustness and weight are negatively correlated (even if this correlation is only weak), then having robust beliefs is a sign of epistemic disvalue. In the absence of a convincing argument why robustness is valuable by itself, it is unclear why a fact-finder's belief lacking robustness should lead to reasonable doubt.

2.2.2 Robustness as the Absence of Exonerating Information

The upshot of the above is that robustness as the resistance of our beliefs to change in the light of new evidence is not a plausible condition for proof of guilt beyond a reasonable doubt. However, when we look at some of the authors who defend a robustness condition, it seems that they have a different (and, I suggest, more plausible) concept in mind. For instance, according Di Bello (2013, 216) the robustness condition is met when the defense had the opportunity to raise challenges to the prosecution's narrative of guilt (and took advantage of this opportunity). He suggests that this typically requires the narrative to be 'complete' – i.e., it should give a detailed account of what happened, which includes, for instance, the perpetrator, the actions that they undertook and their motive for undertaking this action. A complete narrative helps reach robust beliefs because it allows the defense to more easily ask critical questions – e.g., pointing out potential inconsistencies.[3] In other words, making our hypotheses more specific is an important part of meeting the robustness condition. However, this seems to contradict Hamer's point mentioned above that more specific hypotheses tend to be less robust. Does having a more specific hypothesis increase or decrease robustness? The key to resolving this issue is to see that Di Bello is not talking about robustness in the above sense. Whether we critically tested a hypothesis does not matter for how our belief in that hypothesis *would* change if we *were* to receive new evidence. Rather, it relates to whether we could realistically *expect* to find information that would change our belief if we were to search further. This is what critically testing a hypothesis assures us of – that there is no convincing counterargument that we overlooked.

Other recent defenders of a robustness requirement also (implicitly) employ the idea that robustness refers to the (probable) absence of undiscovered exculpatory information. For instance, Mackor & Van Koppen (2021) suggest that robustness is primarily an assessment of the quality of the search for evidence and possible alternative hypotheses. Similarly, Dahlman & Nordgaard (2023) argue that if important evidence was missed during the investigation, then the case for guilt lacks robustness. Neither of these suggestions is about the hypothetical situation in which we actually find new information and how the fact-finder's degree of belief would then change. What a thorough search assures us of is that there is no such (currently undiscovered) evidence of a vast conspiracy. It is this notion of robustness, as a justified belief that no exonerating information was missed, that I defend.

To see how the two notions of robustness may come apart, consider a person who is highly diligent in looking for alternative explanations, facts and arguments against their belief, but who did not find any such information. Their belief is robust in the sense that they are justified in believing that they did not miss any information that contradicts their belief. However, also presume that this person's belief is fickle – if they had found even the slightest scrap of information, they would have immediately revised their belief. Their belief is therefore unrobust in terms of resistance to contradicting information.

---

[3] I explore this idea in more detail in Sect. 5.

3 Reasonable Doubt from Unknown Information

On the interpretation that I propose, both the robustness and evidential probability conditions are about a justified belief that we did not overlook any exculpatory information.[4] However, suppose that we are not very confident that we did not miss anything exculpatory. Why should this lead to a reasonable doubt? In this section I argue that the answer to this question lies in two of the central aims of criminal trials: we want to minimize the number of errors made (*error minimization*) and, to the extent that we do make such errors, we would prefer them to be false acquittals rather than false convictions (*error distribution*). Taking into account the possibility that we may have missed exculpatory information serves these aims.

Before I turn to my own account, I want to briefly consider what I believe to be the strongest argument for why possible overlooked information should *not* lead to a reasonable doubt. This argument will act as the foil for my own position. When we are faced with potential missing evidence, a reasonable response is to gather further information (to the extent that this is cost-effectively possible) (Nance, 2008; 2016). In criminal cases this 'burden of production' will typically fall on the prosecution, unless the defendant has sole access to the missing evidence (Nance, 2008, 278). To the extent that parties fail to meet this burden, sanctions may be imposed on them, especially if the missing information is the consequence of the prosecution or defense acting in a culpable way – for instance if they kept important evidence behind. However, suppose that no further information can be cost-effectively produced. This, we could argue, leaves the fact-finder with no choice other than to make the best use of whatever information they do have (Nance, 2016, 124–137; Biedermann & Vuille, 2019, 17). Even if we know that we likely missed something, such missing information does not give us any reason to change our degree of belief in guilt. After all, we do not know what conclusions the missing information supports. It could be exculpatory but it could also be incriminating. So, according to this argument, if the fact-finder's information indicates that the defendant is probably guilty, then they should convict. To presume that missing information should always benefit the defendant – i.e., to lead to reasonable doubt - is to have, what Laudan (2006, 119) calls a 'pro-defendant bias'. As he points out, the reason why we set the proof standard to a high degree of probability is to distribute the errors fairly. The standard helps us ensure that, to the extent that errors are made, these are mostly false acquittals, not false convictions. However, as Laudan (2006, 119–144) points out, many legal scholars

---

[4] That missing information can be relevant for criminal proof has also been defended by several authors who argue that proof of guilt beyond a reasonable doubt requires that our evidence is 'sufficiently complete,' (e.g., Kaye, 1986; Di Bello, 2013). Roughly speaking this condition entails that the prosecution should present any items of evidence that can be reasonably expected of them. More specifically, no items of evidence should be missing that we would expect to have if the prosecution scenario were true and that the most important parts of the scenario should be backed up with evidence (Di Bello, 2013, 208–210). However, though similar, my proposal is broader than this. Whether we have missed something relates not only to specific, identifiable items of evidence – e.g., particular witnesses not being presented at trial. We may also have reasons to believe that important evidence is missing even if we do not know what this evidence is, for instance if the search for further evidence was lackluster. Furthermore, exculpatory information can, for instance, also take the form of unconceived alternative hypotheses and insights into the connections between the evidence and the hypothesis.

have the intuition that other parts of the criminal proof system should also benefit the defendant – e.g., that deficiencies in the production of evidence always benefit the accused. But, Laudan argues, this intuition is mistaken. It increases the number of false acquittals beyond the ideal distribution, which is already enshrined in the proof standard, thereby ignoring the negative utilities associated with that kind of error. Similarly, one could argue, because we cannot know what the missing information would support, we may not presume that a possibility of overlooked information should necessarily benefit the defendant. Hence, the fact-finder should decide based on their available information in order to achieve the optimal error distribution.

Though the above argument has *prima facie* merit, it ultimately fails. In particular, it relies on what it rejects. The argument presumes that a high proof standard leads to a just error distribution. However, this is not always true – this claim only holds if the fact-finder generally assigns high probabilities of guilt to the guilty and low probabilities of guilt to the innocent. If there is no correlation (or a weak correlation) between the fact-finder's beliefs about guilt and actual guilt, then a requirement of a high degree of belief for conviction will not necessarily lead to errors being shifted in favor of false acquittals. To give a simple (and admittedly silly) example, imagine a fact-finder who convicts by rolling a 10-sided die. If the die lands on 10, they convict. If not, they acquit. There is no reason to presume that those cases in which the fact-finder rolls a 10 will tend to be cases in which the defendant is guilty and those in which they do not roll a 10 will tend to be cases in which the defendant is innocent. There is therefore no reason to expect a just error distribution. Furthermore, such a decision-procedure would also lead to an unacceptably high number of errors, as there would be no necessary correlation between the strength of the evidence and the decision made. This point relates to a second key goal of criminal proof, *error minimization*. We want verdicts to be based on accurate factual beliefs as much as possible, thereby minimizing the number of errors made (see e.g., Nance, 2007, 163; Ho, 2008; Goldman, 2001; Stein, 2005). However, this condition is satisfied only if the fact-finder is able to discriminate between the guilty and the innocent; we want them to be a reliable assessor of who is guilty and who is not.

What is required for the fact-finder's judgments to be reliable? As Laudan (2006, 73) points out, whether a high proof standard leads to the correct error distribution depends in part on the completeness of the evidence and the validity of the fact-finder's inferences from that evidence (Laudan, 2006, 73).[5] In other words, does our evidence contain the most important facts and does the fact-finder assign the appropriate strength to this evidence? It is not difficult to read the criteria of robustness and evidential probability into this remark. To give a simple example of the underlying idea, consider a fact-finder who always has to draw conclusions from a very incomplete case file and who consistently interprets the evidence that they do have in unreasonable ways. Obviously, we would not expect such a fact-finder to discriminate well between the guilty and the innocent and hence, we would not expect their decisions to yield a just error distribution.

---

[5] It also relies on the distribution of truly innocent and truly guilty defendants who go to trial (Laudan, 2006, 75).

Apart from error distribution, many authors point out that robustness and evidential probability are also key requirements for *error minimization*. To begin with the notion of evidential probability, Laudan (2006, 79) argues that the beyond a reasonable doubt standard should not be purely a measure of personal belief, because a purely subjective standard has no necessary connection with error distribution. Such a standard allows for conviction based on weak sets of evidence for guilt acquittal in cases where there is strong evidence for guilt as it means that almost any belief can count as rational, even when it does not align with what the evidence supports. Similarly, Allen and Pardo (2019, 9–10) argue that unconstrained subjective beliefs have "no necessary relationship to advancing accurate outcomes" as they could be any number at all and need not be constrained by the evidence. Evidential probabilities ensure that rational degrees of belief *are* constrained by the evidence. Nance (2008, 270) offers a similar reason for why we are after evidential probabilities in legal proof. He suggests that this is to ensure that the relevant probability is "both well-considered and productive of accurate verdicts."

The requirement of robustness (understood as the probable absence of overlooked exculpatory information) similarly contributes to the reliability of criminal fact-finding. To illustrate, imagine a fact-finder who always assigns the correct strength to the available evidence. However, imagine that their set of evidence is consistently the result of a sloppy investigation, which overlooks most of the relevant facts. We would not expect this fact-finder's judgments to be especially accurate, as there is no reason to presume that this evidence is indicative of the actual status of the hypothesis that the defendant is guilty. So, as Ho (2008, 167) suggests "if the trier of fact is aware that the available evidence adduced in support of a hypothesis is significantly incomplete, that too much of relevance is as yet hidden from her, that 'there is a significant chance that there is a better explanation' for the event in question, she would not be justified in believing that the hypothesis is true."

In the quote above, Ho mentions the possibility of unconceived alternatives as one of the sources of reasonable doubt. This remark connects to a point made by various Bayesian philosophers of science. As they point out, we are only justified in assigning a probability to any given hypothesis if we presume that the set of hypotheses that we consider exhausts the probability space. For instance, Salmon (1990) proposes a Bayesian account which does not presume that our set of hypotheses is exhaustive, but on which we only consider our conceived alternatives when evaluating the confirmation of a given hypothesis. But, as Rowbottom (2016, 3) points out in response, if we only consider conceived alternatives, this means letting go of the assumption that we are evaluating whether a theory is "truth-like." Similarly, Wenmackers & Romeijn (2016) propose an 'open minded' version of Bayesianism, which drops the assumption "implicit in standard Bayesianism – that the correct empirical hypothesis is among the ones currently under consideration." However, they admit that their approach "fails to provide us with the required normative guidance" about the absolute confirmation of scientific theories, because it only tells an agent what to believe if she supposes "that the true theory is among those currently under consideration" (Wenmackers & Romeijn, 2016, 1243). In other words, unless we presume that our set of hypotheses is exhaustive, we must let go of the assumption that the probability that we assign to a hypothesis is indicative of the actual truth-value of that hypothesis

(Jellema, 2022). What the above discussion suggests is that the same is true of other kinds of missing information. If we cannot presume that we have not missed anything exculpatory, then we also cannot presume that a high degree of belief in guilt is an indicator that the defendant is actually guilty (or, conversely, that a low degree of belief in guilt indicates that the defendant is innocent).

So, to summarize the above, we want robust evidentially calibrated beliefs for the sake of achieving a just error distribution and for minimizing the number of factual errors. If a fact-finder were to act upon a belief while they know that they have likely missed a great deal of relevant information, then the risk of error is too great. And, in high-stakes situations, which criminal cases regularly are, we arguably want to err on the side of caution by opting for the less risky option (Horowitz, 2014; Henderson, 2021, 6–7). Because a false acquittal is less costly than a false conviction, the fact-finder should acquit if we cannot justifiably believe that we did not overlook any exculpatory information. Whether we are justified in presuming this will depend on our 'higher-order evidence'. I now turn to the meaning of this notion, discuss how our resulting higher-order beliefs (expressed as 'higher-order probabilities') relate to first-order beliefs and offer some examples of higher-order evidence.

4 Higher-Order Probability and Higher-Order Evidence

Information about whether we missed anything is not evidence in the ordinary sense, as it does not relate directly to the hypothesis under consideration (that the defendant committed the alleged acts). The reason for this is that such information does not sanction a specific change in our degree of belief in the hypothesis. After all, we do not know what the missing information would support. As Hamer (2012, 136) writes, "probabilistically, it is not possible to take account of unavailable [information] for the simple reason that it is unavailable and its content is unknown." Instead I want to suggest that we can better think of it as 'higher-order evidence'. Higher-order evidence is a well-known concept from epistemology. There are various ways of spelling out this idea. For instance, Henderson (2021) mentions several characterizations that epistemologists have given of the concept, including 'evidence concerning the reliability of our own thinking about some particular matter' (Christensen, 2016), 'evidence about what your evidence supports' (Sliwa & Horowitz, 2015) and evidence that 'induces doubts that one's doxastic state is the result of a flawed process' (Lasonen-Aarnio, 2014). So, roughly speaking, higher-order evidence is evidence about how reliable the conclusions are that we drew from our first-order evidence. Though the question of missing evidence has, as far as I am aware, not been linked to the idea of higher order evidence, the latter concept fits closely with what I suggested in the previous section, namely that the possibility of missed exculpatory information can induce doubt about the reliability of the fact-finder's inferences from that evidence.

In this section I distinguish several types of higher-order evidence, which may jointly justify the belief that we did not miss any exculpatory information. However, I first want to discuss how such higher-order evidence impacts first-order beliefs.

4.1 Higher-Order and First-Order Belief

As with any form of evidential reasoning, our higher-order evidence can be stronger or weaker. So, how justified we are in presuming that we have not missed anything comes in degrees. Hence, we could understand my proposal in terms of 'higher-order probabilities' – which would be the probability that we have missed exculpatory information. There are various ways of interpreting the notion of a higher-order probability, but not all of them are suitable for legal proof. For instance, on one interpretation a higher-order probability expresses the rate at which the fact-finder tends to make errors in their first-order judgments. The higher this error rate, the lower the higher-order probability. Such an error rate can be taken into account within the first-order belief through 'calibration', where we lower the first-order probability in proportion to the error rate (Schoenfield, 2015). Malcai & Rivlin (2021) suggest that we can use this idea of higher-order probability in legal proof. For instance, they propose that sometimes a judge might lower their credence in the guilt of a defendant if they know that they have erroneously convicted innocent defendants in the past. However, this suggestion strikes me as problematic for several reasons. First, this only works for fact-finders who have a track record – i.e., judges, not juries. But even in the case of judges, we rarely, if ever, have reliable data about their error rate. Second, historical data is a poor guide to one's error rate. Each case is unique and simply because errors were made in the past does not mean that errors will be made at the same rate in future cases. The extent to which we would expect errors will depend on the particulars of the case at hand. Finally, my account of higher-order probability relates to the possibility of overlooking information. As said, evidence of having overlooked something does not sanction any particular change in our first-order degrees of belief.

I want to propose that we should instead understand the relationship between our first- and higher-order beliefs in terms of a full-belief framework – where we either believe something or not. Within such a framework, higher-order evidence is usually conceptualized as an undercutting defeater (Lasonen-Aarnio, 2014; Henderson, 2021). An undercutting defeater casts doubt on the connection between the first-order evidence and the belief concerning the first-order proposition. As said, we can think of our higher-order belief (about whether we are justified in our first-order belief) as coming in degrees. But these higher-order probabilities do not change the level of our first-order credences. Rather, when our higher-order probability is too low, this severs the link between the fact-finder's belief that the defendant is probably guilty and the conclusion that the evidence supports the defendant's guilt. In other words, the fact-finder is not justified in assigning a (high) first-order probability to the defendant's guilt if their higher-order evidence does not justify them in believing that their set of information is (sufficiently) exhaustive. The requirements for proof of guilt beyond a reasonable doubt are then not met.

4.2 Types of Higher-Order Evidence

Higher-order evidence may lead to doubt about first-order judgments of the defendant's guilt. Different types of such evidence may occur in criminal proof. One important kind of higher-order evidence is whether the fact-finder had the opportunity to

consider the evidence carefully. To give a fictional example, at the start of the 1957 film 12 Angry Men, only one of the jurors, Davis, votes not guilty. Another juror asks him whether he believes the defendant's story. Davis replies: "I don't know whether I believe it or not. Maybe I don't. (…) There were eleven votes for guilty. It's not so easy for me to raise my hand and send a boy off to die without talking about it first." What the juror indicates is that the issue is not whether or not he is convinced at that specific moment, but about carefully considering the evidence and potential explanations for this evidence before arriving at a verdict.

Another, related type of higher-order evidence is discussed by Di Bello (2013, ch. 7.5). He argues that robustness depends in part on the degree to which the defense had an opportunity to level charges against the prosecution's case and the degree to which they took advantage of this opportunity. In many countries there are specific legal guarantees that ensure that defendants have enough monetary, legal, intellectual, and evidentiary resources to exercise their right to a defense, such as the right to effective council. If such procedural rights are not respected, then this alone may lead to the case being dropped or to an acquittal. Additionally, such a violation of rights may lessen the degree to which the fact-finder believes that they understand what the evidence actually supports and whether nothing was overlooked.

A third type of higher order evidence is evidence about the quality of the underlying investigation. All other things being equal, the better the search for evidence, alternative hypotheses, connections between the evidence and weaknesses in the case, the more reason we have to presume that our set of information is good enough to base a reliable belief on. The quality of such a search depends in part on the amount of time and resources spent on it as well as on the imaginative faculties of the investigators. It also depends on epistemic virtues on the part of investigators, such as open-mindedness and perseverance as well as on how methodically they construed their investigation. Whether investigators displayed these virtues can also depend on the nature of the case. For instance, as Amaya (2015, 517) points out, in emotionally disturbing cases, investigators may be more likely to be biased, thereby failing to conceive of plausible alternatives. Additionally, we may have information that important evidence was not collected during the investigation (Dahlman & Nordgaard, 2023). For instance, an important witness may not have been heard.

The quantity and quality of our set of information also matters in several ways. We may firstly have too much information to adequately make sense of it. For instance, the case may be accompanied by a thick case file, indicating a great deal of potentially relevant evidence. It may then be difficult to see how all the facts hang together. Additionally, having too much information can create 'noise', where irrelevant information drowns out the (more) relevant. The same goes for having too many possible explanations or having to make sense of many competing arguments. Such an abundance of information can make it difficult to judge whether all plausible alternatives were considered and whether all relevant evidence was collected. Second, some evidence, such as statistical information, is known to be difficult to interpret for lawyers (Malcai & Rivlin, 2021, 29). This can increase the chance that relevant arguments are overlooked, as it is more difficult to come up with such arguments, thereby decreasing the fact-finder's confidence in their own probability assessment.

Another type of higher-order evidence is the quality of our conceived hypotheses. If none of our current alternatives explain the evidence well, then we have good reasons to suspect that either there is a better explanation that we have not conceived of, or that some of the evidence is misleading. For example, consider a situation in which none of our current scenarios explain the testimony of multiple witnesses well. This might make us think that we have overlooked a scenario that *does* explain these testimonies adequately, or that we have failed to realize that one of the witnesses is lying. Also, as Dahlman & Nordgaard (2023) point out, the better the case for guilt, the less likely it is that it will be overturned by newly discovered evidence. If the case for guilt only barely meets the threshold for conviction, then it is more easily overturned.

A final way in which our existing explanations can be of a high quality is that they are sufficiently detailed. It is to this form of higher-order evidence that I turn now.

5 Explanation-Based Thinking, Specificity and Robust Evidential Probability

If the hypotheses under consideration during the investigation and subsequent trial were well-specified, this arguably increases the degree to which the fact-finder is justified in assuming that their belief in the defendant's guilt is robust and evidentially calibrated. The reason why I single out this particular type of evidence is that it leads to a novel answer on how the Bayesian and explanation-based accounts of rational legal proof can be combined. More and more evidence scholars suggest that these two accounts are compatible and may also complement one another (Mackor, Jellema & van Koppen, 2021). However, discussions about the relationship between the two mainly focus on how explanation-based thinking can lead to a justified high credence in first-order propositions (e.g., Hedden & Colyvan, 2019; Biedermann & Vuille, 2019, 18–20; Gelbach, 2019, 169; Welch, 2020). What I want to suggest here is that explanation-based thinking can (also) complement the Bayesian account on the level of the fact-finder's higher-order belief.

Explanation-based accounts cast rational legal proof in terms of competing explanations of the evidence. On this account, whether guilt is proven depends on the plausibility of the available guilt and innocence explanations. As most legal explanationists use the term, 'plausibility' refers to the extent that an explanation exhibits 'explanatory virtues', such as internal coherence and fit with background beliefs about the world. One explanatory value that is sometimes mentioned, both by explanationists in the law and in the philosophy of science is 'specificity' (or 'preciseness') (Thagard, 1978; Pennington & Hastie, 1991; Ylikoski & Kuorikoski, 2010). The virtue of specificity is especially important in criminal proof, where the relevant explanations often take the form of scenarios – narratives that describe a sequence of events which led, for instance, to the death of the victim (Allen & Pardo, 2019, 13, n86; Mackor & Van Koppen, 2021). We typically want our scenarios to be sufficiently detailed. For instance, according to Bennett & Feldman (1981) they should ideally contain a central action and describe a context that make this action understandable, in the form of a description of the scene, a motive, a central actor and resulting consequences. According to Pennington & Hastie (1993) a complete scenario includes an initiating

event, a psychological response to this event, a goal, a resulting action, and consequences. Apart from these details the scenarios under consideration should arguably be sufficiently specific about the time and place of the alleged events as well as about how those events took place. However, details such as motive, time and place will typically not be a part of the legal definition of the alleged criminal offense. So why do fact-finders consider detailed hypotheses to be more convincing?

The idea that more detailed hypotheses are preferable to general ones raises a problem for the explanation-based accounts. This account shares the idea with Bayesianism that we only want to accept conclusions that are (very) probably true (Allen & Pardo, 2018, 1580; 2019, 17). So, we should only accept an explanation as true if it is very probably correct. However, there is a tension between this aim and the wish to have detailed hypotheses. All else being equal, more detailed hypotheses are less probable than less detailed ones. For example, consider two explanations of why a person died:

> The defendant killed the victim.
> The defendant killed the victim with a hammer, around 5 p.m., after they got in
> a fight over an unpaid loan.

The second hypothesis can never be more probable than the first, as it is a more specific version of the former. Whenever the second is true, so is the first but there are situations where the first is true while the second is not. So, if our aim is to accept only highly probable explanations, why would we want to reason in terms of the latter rather than the former? More generally, on the explanation-based account we reason in terms of a small set of non-exhaustive, specific hypotheses (Allen & Pardo, 2019, 11−2). But why not take the full probability space into consideration? And why would.

This worry brings us back to the argument by Di Bello (2013) that more specific hypotheses make our beliefs more robust, where robustness means the probable absence of exonerating information. This is an interesting argument that has, as far as I am aware, not received much attention in legal proof scholarship. Di Bello's point is that specific scenarios can more easily be subjected to scrutiny as it is easier to ask critical questions of specific details about broad hypotheses. Because it is easier to scrutinize specific hypotheses there is less chance that a flaw in that hypothesis escaped our attention.

Well-specified scenarios also have other benefits. For instance, as Pennington & Hastie (1993) argue, thinking in terms of such scenarios helps fact-finders make sense of complex sets of evidence. One way in which scenario-based thinking helps us make sense of the evidence is that it helps reason about which facts are relevant in the given case. For example, imagine a case in which there is a medical report which states that the defendant has a limp. Such a report is not, by itself, evidence of anything. It is simply one of the countless facts regarding the defendant. However, suppose that the report is evidence in a burglary case. The prosecution scenario in this case may stipulate that the defendant entered the house through the garden. As a result, the report may become pertinent information. For instance, we may then want to look for footprints in the garden and see whether the pattern of these prints

indicated that the perpetrator had a limp. If the footprints match someone with a limp, the report supports the guilt of the defendant. So, because the scenario specified how the perpetrator entered the house, investigators could consider what evidence was pertinent. Conversely, if this had not been specified, both the footprints and the fact that these prints matched the limp of the defendant might have been missed.

A sufficiently specific scenario can also help us discover internal contradictions within that scenario. For instance, suppose that the prosecution's scenario entails that the limping defendant fled on foot but that he could not have traveled the required distance within the allotted time frame given this limp. The discovery of this difficulty required a detailed scenario which stipulates which path was taken within what time-frame.

Another benefit of thinking in terms of well-specified scenarios is that it can lead us to consider alternative scenarios. For example, suppose that we search for further footprints in the garden but do not find any. This can make us wonder whether there is an alternative explanation for why no footprints were found there. For instance, the defendant might not have entered through the yard.[6] To come up with such alternative scenarios it is helpful to know which part of our conceived scenario(s) is implausible, as we may be able to think of a similar scenario without this implausible element.

The above remarks are only some of the ways in which comparing well-specified scenarios help us make sense of the evidence. The upshot of this is that when we think in terms of multiple detailed scenarios, this makes it easier to determine whether we may have missed something. If nothing is found, this strongly suggests (all else being equal) that the fact-finder's degree of belief is robust and evidentially calibrated. What we sacrifice in first-order probability when we make our hypotheses more specific is then offset by a gain in higher-order probability.

That explanation-based thinking helps make sense of the evidence is a well-known idea. For instance, Nance (2016, 84) observes that "[o]ne main motivating concern of those who press the explanatory approach is that [probabilistic accounts] focus on the end product of deliberation, rather than the process of arriving there, giving no direction to jurors as to how to go about assessing the evidence in the case." What is new in my analysis is the connection between this idea and the notions of robustness, evidential probability, higher-order evidence and higher-order probability.

6 Conclusion

In criminal trials we are after proof of guilt beyond a reasonable doubt. Many agree that this at least requires the fact-finder to believe that the defendant very probably committed the alleged acts. However, a high degree of belief alone is insufficient for meeting this standard. After all, this belief may be unreasonable. In this article I focused on two popular suggestions for additional criteria: that the fact-finder's belief should be an evidential probability and that it should be robust. As I interpreted these

---

[6] In other words, we generate a new scenario by asking 'how else could we explain this surprising fact'? This is a well-known advantage of abductive reasoning (i.e. explanation-based thinking) over Bayesianism, which says nothing about the generation of new hypotheses (Lipton, 2004, 116).

terms, both are about the possibility of overlooked exculpatory information. I argued that if we cannot justifiably presume that there is no such overlooked information, then there should be reasonable doubt. Whether this presumption is justified or not will depend on our higher-order evidence. I surveyed several kinds of higher-order criminal evidence. I ended this article by looking at a particular type of higher-order evidence, namely the specificity of our hypotheses. This led to a novel way of combining the Bayesian and explanation-based account of rational legal proof.

Declarations

Competing Interests The author has no competing interests to declare.

References

Acree, M. C. (2021). *The Myth of Statistical Inference*. Springer.

Allen, R. J. (2017). The nature of juridical proof: Probability as a tool in plausible reasoning. *The International Journal of Evidence & Proof*, *21*(1–2), 133–142.

Allen, R. J., & Pardo, M. S. (2007). Probability, explanation and inference: A reply. *The International Journal of Evidence & Proof*, *11*(4), 307–317.

Allen, R. J., & Pardo, M. S. (2019). Relative plausibility and its critics. *The International Journal of Evidence & Proof*, *23*(1–2), 5–59.

Amaya, A. (2015). *The tapestry of reason: An inquiry into the nature of coherence and its role in legal argument*. Hart and Portland.

Bennett, W. L., & Feldman, M. S. (1981). *Reconstructing reality in the courtroom: Justice and judgment in American culture*. Quid Pro Books.

Biedermann, A., & Vuille, J. (2019). The decisional nature of probability and plausibility assessments in juridical evidence and proof. *International Commentary on Evidence*, *16*(1), 1–30.

Dahlman, C., & Nordgaard, A. (2023). Information economics in the criminal standard of proof. *Law, Probability and Risk*.

Dahlman, C., Wahlberg, L., & Sarwar, F. (2015). Robust trust in expert testimony. *Humana Mente*, *28*, 17–37.

Di Bello, M. (2013). *Statistics and probability in criminal trials*. Stanford University.

Gardiner, G. (2019). The Reasonable and the Relevant: Legal Standards of Proof. *Philosophy & Public Affairs*, *47*(3), 288–318.

Gelbach, J. B. (2019). It's all relative: Explanationism and probabilistic evidence theory. *The International Journal of Evidence & Proof*, *23*(1–2), 168–175.

Goldman, A. I. (2001). Quasi-objective Bayesianism and legal evidence. *Jurimetrics*, *42*, 237–260.

Hacking, I. (2001). *An introduction to probability and inductive logic*. Cambridge university press.

Hamer, D. (2012). Probability, anti-resilience, and the weight of expectation. *Law, Probability and Risk*, *11*(2–3), 135–158.

Hedden, B., & Colyvan, M. (2019). Legal probabilism: A qualified defence. *Journal of Political Philosophy*, *27*(4), 448–468.

Henderson, L. (2021). Higher-order evidence and losing one's conviction. *Noûs*.

Ho, H. L. (2008). *A philosophy of evidence law: Justice in the search for truth*. Oxford University Press on Demand.

Horowitz, S. (2014). Epistemic akrasia. *Noûs*, *48*(4), 718–744.

Horwich, P. (1982). *Probability and evidence*. Cambridge University Press.

Hunt, I., & Mostyn, J. (2020). Probability reasoning in judicial fact-finding. *The International Journal of Evidence & Proof*, *24*(1), 75–94.

Jellema, H. (2022). Reasonable doubt from unconceived alternatives. Erkenntnis.

Kaye, D. H. (1986). Do we need a calculus of weight to understand proof beyond a reasonable doubt. *Boston University Law Review*, *66*, 657–672.

Keynes, J. M. (1921). *A treatise on probability*. Macmillan and Company, limited.

Lasonen-Aarnio, M. (2014). Higher-order evidence and the limits of defeat. *Philosophy and Phenomenological Research*, *88*(2), 314–345.

Laudan, L. (2006). *Truth, error, and criminal law: An essay in legal epistemology*. Cambridge University Press.

Lipton, P. (2004). *Inference to the best explanation*. Routledge.

Logue, J. (1997). Resiliency, robustness and rationality of probability judgements. *International Studies in the Philosophy of Science*, *11*(1), 21–34.

Mackor, A. R., & van Koppen, P. J. (2021). The Scenario Theory about Evidence in Criminal Law. In C. Dahlman, A. Stein, & G. Tuzet (Eds.), *Philosophical Foundations of Evidence Law* (pp. 215–230). Oxford University Press.

Mackor, A. R., Jellema, H., & Van Koppen, P. J. (2021). Explanation-based approaches to reasoning about evidence and proof in criminal trials. In Brozek, Bartosz & Hage, Jaap (Eds.), *Law and Mind: A Survey of Law and the Cognitive Sciences* (pp. 431–469). Cambridge University Press.

Malcai, O., & Rivlin, R. (2021). Reasonable Self-doubt. *Criminal Law and Philosophy*, *15*(1), 25–45.

Nance, D. A. (2007). Allocating the Risk of Error: Its Role in the Theory of Evidence Law. *Legal Theory*, *13*(2), 129–164.

Nance, D. A. (2008). The weights of evidence. *Episteme*, *5*(3), 267–281.

Nance, D. A. (2016). *The Burdens of Proof*. Cambridge University Press.

Pennington, N., & Hastie, R. (1991). A Cognitive Theory of Juror Decision Making: The Story Model. *Cardozo Law Review*, *13*(2–3), 519–558.

Pennington, N., & Hastie, R. (1993). Explaining the evidence: Tests of the Story Model for juror decision making. *Journal of Personality and Social Psychology*, *62*(2), 189–206.

Redmayne, M. (2003). Objective probability and the assessment of evidence. *Law, Probability and Risk*, *2*(4), 275–294.

Risinger, D. M. (2013). Reservations about likelihood ratios (and some other aspects of forensic 'Bayesianism'). *Law, Probability and Risk*, *12*(1), 63–73.

Salmon, W. (1990). Rationality and objectivity in science or Tom Kuhn meets Tom Bayes. In C. Wade Savage (Ed.), *Scientific Theories* (Vol. 14, pp. 175–204). Oxford University Press.

Schoenfield, M. (2015). A dilemma for calibrationism. *Philosophy and Phenomenological Research*, *91*(2), 425–455.

Skyrms, B. (1977). Resiliency, propensities, and causal necessity. *The Journal of Philosophy*, *74*(11), 704–713.

Sliwa, P., & Horowitz, S. (2015). Respecting all the evidence. *Philosophical Studies*, *172*(11), 2835–2858.

Spottswood, M. (2019). On the limitations of a unitary model of the proof process. *The International Journal of Evidence & Proof*, *23*(1–2), 75–81.

Stein, A. (2005). *Foundations of Evidence Law*. Oxford University Press.

Thagard, P. R. (1978). The best explanation: Criteria for theory choice. *The Journal of Philosophy*, *75*(2), 76–92.

Urbaniak, R. (2018). Narration in judiciary fact-finding: A probabilistic explication. *Artificial Intelligence and Law*, 1–32.

Welch, J. R. (2020). Rebooting the new evidence scholarship. *The International Journal of Evidence & Proof*, *24*(4), 351–373.

Wenmackers, S., & Romeijn, J.-W. (2016). New theory about old evidence. *Synthese*, *193*(4), 1225–1250.

Williamson, T. (2002). *Knowledge and its Limits*. Oxford University Press on Demand.

Wittlin, M. (2019). Common problems of plausibility and probabilism. *The International Journal of Evidence & Proof*, *23*, 184–190.

Ylikoski, P., & Kuorikoski, J. (2010). Dissecting explanatory power. *Philosophical Studies*, *148*(2), 201–219.