

Evidence, Risk, and Proof Paradoxes: Pessimism about the Epistemic Project¹

Giada Fratantonio University of Edinburgh

Forthcoming in *The International Journal of Evidence and Proof*

[this is a final draft. Please cite published version when possible]

Abstract

Why can testimony alone be enough for findings of liability? Why statistical evidence alone can't? These questions underpin the "Proof Paradox" (Redmayne 2008, Enoch et al. 2012). Many epistemologists have attempted to explain this paradox from a purely epistemic perspective. I call it the "Epistemic Project". In this paper, I take a step back from this recent trend. Stemming from considerations about the nature and role of standards of proof, I define three requirements that any successful account in line with the Epistemic Project should meet. I then consider three recent epistemic accounts on which the standard is met when the evidence rules out modal risk (Pritchard 2018), normic risk (Ebert et al. 2020), or relevant alternatives (Gardiner 2019 2020). I argue that none of these accounts meets all the requirements. Finally, I offer reasons to be pessimistic about the prospects of having a successful epistemic explanation of the paradox. I suggest the discussion on the proof paradox would benefit from undergoing a "value-turn".

Keywords: Evidence-law; proof paradoxes; Epistemology; Legal Risk; Relevant Alternatives; modal risk; normic Risk.

1. Introduction

According to a traditional interpretation of the civil standard of proof, i.e., *preponderance of evidence*, the standard of proof is met when the available evidence makes it more likely than not that the defendant is liable. But consider the following two cases:

¹ I have presented parts of this paper in 2019 and 2020 at the Legal Philosophy Workshop at the Surrey Centre for Law and Philosophy, at the Responsibility Knowledge and Belief Workshop hosted by the ERC-funded Roots of Responsibility Project at UCL, at the Spring Workshop hosted by the ERC-funded Competence and Success in Epistemology and Beyond Project at the University of Helsinki, at Maastricht school of Law, and at the Edinburgh Legal Theory Group. I am grateful to everyone who gave me feedback on those occasions and, in particular, to Aness Webster who gave a response to my talk in Surrey. Further thanks to Dylan Balfour, Daniel Baras, Daniel Drucker, Philip Ebert, Antonella Frasca Caccia, Jaakko Hirvelä, Maria Lasonen-Aarnio, Lilith Newton, Angie O' Sullivan, Petronella Randell, Xington Wei, and Di Yang for useful discussion on this paper. I am also very grateful to Antony Duff, Claire Field, Lucas Miotto, Dario Mortini, and Martin Smith for written comments on previous drafts of this paper. Finally, special thanks to Martin Smith for countless moments of discussions on evidence, normic support, and proof paradoxes. The work for this paper was supported by the European Research Council (European Union's Horizon 2020 research and innovation program, 758539) and the Arts and Humanities Research Council ("Varieties of Risk," AH/T002638/1).

BLUE BUS

Mr. Brown is run over by a bus on Montgomery Street; 90% of the buses travelling on this street are owned by the Blue Bus Company, and 10% by the Red Bus Company. Mr. Brown, however, couldn't see the colour of the bus. He decides to sue the Blue Bus Company. However, given the only evidence available is statistical evidence, the Blue Bus Company is not and should not be found liable.²

BLUE BUS TESTIMONY

Mr. Brown is run over by a bus on Montgomery Street. He couldn't see which bus hit him. However, a bystander testifies that she saw a blue bus hitting Mr. Brown. No further evidence is presented against the reliability of the person testifying. The expected reliability of the eyewitness testimony is 70%. Given eyewitness testimony is the only evidence available, the Blue Bus Company is and should be found liable.³

In both cases, it's more likely than not that a Blue Bus hit Mr. Brown. And yet, while it seems wrong to impose liability on the basis of numbers alone (as in *BLUE BUS*), there seems nothing controversial in finding the Blue Bus Company liable on the basis of testimony alone (provided the testimony has passed cross-examination and that there's no evidence against her reliability or sincerity)⁴ (as in *BLUE BUS TESTIMONY*). But, as specified in *BLUE BUS TESTIMONY*, the estimated reliability of an eyewitness testimony is way lower than 90%. This is known as the "Proof Paradox" (e.g., Enoch et al. 2012, Redmayne 2008). Given its similarities with notorious epistemic paradoxes, many philosophers have tried to solve this paradox from a purely epistemic perspective. By offering "epistemic accounts" of the proof paradox, they've identified a non-probabilistic "epistemic quality" that: i) non-individualised evidence lacks (e.g., paradigmatic cases of statistical evidence like base-rates); individualised evidence has (e.g., paradigmatic cases of testimony); and that should be required for meeting the different standards of proof. I call this the Epistemic Project.

This paper doesn't aim to solve the Proof Paradox. Instead, after clarifying the nature and scope of the Epistemic Project (§2), it identifies three requirements that every epistemic account should meet: a Value, a Functionalist, and a Feasibility Requirement (§3). The first follows from the aim of the Epistemic Project itself; the other two stem from considerations about the role standards of proof play in managing legal risk. In §4, I consider three recent epistemic accounts which tackle the proof paradox by considering the notion and distribution of legal risk. On these accounts, managing legal risk doesn't require minimising error. It requires ruling out modal risk (Pritchard 2016 2018), normic risk (Ebert et al. 2019), or all the relevant alternatives to the litigated claim (Gardiner 2019 2020). Crucially, I argue that they all

² Redmayne (2008). This is based on the real case *Smith v Rapid Transit Inc.*, 317 Mass. 469, 470, 58 N.E 2d 754, 755 (1945). One might question whether the cases are really analogous (Cf Spottswood forthcoming). Another case generally discussed in this literature is *United States v. Shonubi*, 895 F. Supp. 460 (E.D.N.Y. 1995).

³ E.g., Smith (2016: section 1)

⁴ Henceforth, I'll assume this caveat when talking about a verdict based on testimony "alone".

struggle to meet these requirements. Finally, in §5, I provide reasons to believe other epistemic accounts to be similarly unsuccessful: the Epistemic Project seems to rest on a problematic methodology.

2. The Epistemic Project

2.1. Clarifying the Project

Why did this legal paradox spark particular interest amongst epistemologists? Consider the following notorious case:

LOTTERY

You own a ticket in a fair one-million ticket lottery; your ticket has been drawn but you haven't checked the results yet. On the basis of the odds involved, you believe that your lottery ticket is a loser. Although your ticket is indeed a loser, your belief is not justified. Assuming knowledge entails justification, you cannot know that your ticket is a loser merely on the basis of statistical evidence.⁵

There are structural similarities between *LOTTERY* and *BLUE BUS*. They both involve a proposition made highly probable on the basis of statistical evidence. They generate similar reactions: we are reluctant to have findings of liability on the basis of statistical evidence only, a phenomenon known as the “Wells’ effect” (Wells 1992); we are reluctant to ascribe justification and knowledge to beliefs in lottery propositions. Furthermore, in both cases, our uneasiness seems to go away once we are presented with testimony, rather than statistical evidence, as described in *BLUE BUS TESTIMONY* and in the following variation of *Lottery*:

LOTTERY NEWSPAPER

You own a ticket in a fair one-million ticket lottery. You believe on the basis of a newspaper’s report that your lottery ticket is the winner. The local newspaper is, we can assume, roughly 70% reliable. In this case, it’s plausible to ascribe justification (and, assuming your lottery ticket has in fact won, even knowledge) to your belief that your lottery ticket is the winner.

The *BLUE BUS/BLUE BUS TESTIMONY* pair presents a challenge to a view, often called Legal Probabilism, on which standards of proof should be always understood in terms of probability (Laudan 2006: Ch.3, Hedden and Colyvan 2019). Similarly, in epistemology, *LOTTERY/LOTTERY NEWSPAPER* are used to challenge probabilistic accounts of justification, e.g., the Threshold View, on which one has justification to believe a proposition *p* if and only if one’s evidence makes *p* sufficiently likely (See Achinstein 2001). But what’s so paradoxical about these cases? Roughly put, the paradox arises

⁵ See for example Kyburg (1961); Nelkin (2000), Hawthorne (2004).

because maximising accuracy when delivering a verdict or when deciding what to believe seems the reasonable thing to do, and both Legal Probabilism and the Threshold View serve this aim very well. And yet, we need to explain why we are willing to ascribe justification and have an affirmative verdict in LOTTERY NEWSPAPER and BLUE BUS TESTIMONY respectively, while we are reluctant to do so in LOTTERY and BLUE BUS, which, by stipulation of the case, involve evidence that better serves the aim of accuracy. Ever since, epistemologists have offered various solutions to the lottery paradox. Given the striking similarities between the two paradoxes, it's no surprise that one might think that a purely epistemic explanation is available for the legal puzzle as well. This parallel between the two paradoxes underpins what I call the "Epistemic Project" of the Proof Paradox. The Epistemic Project rests on the idea that standards of proof should be understood in terms of some non-probabilistic epistemic feature that the available admissible evidence should have to satisfy such standard. In other words, the Epistemic Project identifies the "epistemic quality" that non-individualised evidence lacks, individualised evidence has, and that is necessary for meeting the standard of proof.^{6,7} According to what I call the "Doxastic Approach", meeting the standards of proof requires evidence that generates a specific *doxastic state* towards the proposition that the defendant is liable or guilty: full belief (Roth 2010; Buchak 2014), knowledge (Littlejohn, 2017), or probabilistic knowledge (Moss 2018). Alternatively, one could pursue the "Propositional Approach", which focuses on the relation between the admissible evidence and the proposition that the defendant is liable. For instance, it's been argued that evidence should be sensitive (Enoch et al. 2012)⁸ or causally connected to the litigated claim (Thompson, 1986); make it probable that we know the litigated claim (Blome-Tillman, 2017); rule out relevant alternatives where the litigated claim is false (Gardiner 2019); rule out *normic* or *modal* risk that the litigated claim is false (Ebert et. al 2020; Smith 2016 2018; Pritchard 2016 2018). In this paper, I take a step back from this recent trend and argue that the Epistemic Project is wrongheaded. However, before doing that, I will disambiguate the different ways of developing the Epistemic Project, and provide preliminary reasons to doubt the plausibility each of them.

2.2. Preliminary reasons to be suspicious of the Epistemic Project

First, we can distinguish three epistemic "sub-projects":

- A *descriptive* project that describes how the law operates;

⁶ Henceforth, when I talk about "epistemic accounts" and "Epistemic Project", I mean those accounts that identify an epistemic condition *other than/instead of* probability.

⁷ Note that the Proof Paradox isn't meant to question whether naked statistical evidence is *admissible*, but whether it's *sufficient* for meeting the standard of proof.

⁸ According to Enoch et al. (2012), however, sensitivity is required for *instrumental* reasons.

- An *explanatory* project that explains lay people’s intuitions;
- A *normative* project that defines how the law should operate.⁹

The epistemic literature on this topic has been mainly concerned with the explanatory and the normative project. A classic paper on this topic would start off by taking the intuitions generated by the proof and lottery paradoxes at face value; it would try to explain them in a unified fashion, and it would conclude with a *normative* claim on how standards of proof *should* be understood (For surveys: Redmayne 2008, Ross 2020, Gardiner forthcoming).

Second, note that, while taking BLUE BUS/BLUE BUS TESTIMONY as the paradigmatic version of the Proof Paradox, epistemologists have often generalised their conclusions to other cases: they have tried to provide a *unified* epistemic account that is sensitive to the different standards of proof¹⁰, and that can also be applied to versions of the paradox involving other kinds of individualised and non-individualised evidence (e.g., CCTV footage and cold-hit DNA evidence¹¹ respectively). In other words, the Epistemic Project aims to identify a *general* epistemic inadequacy characterising non-individualised evidence, which also explains the epistemic adequacy of individualised evidence. This spirit has been recently made explicit by Sarah Moss:

“The problem of statistical evidence should not be explained by any feature of the reasonable doubt standard in particular, but rather by some feature that all standards of proof have in common. In other words, the problem of statistical evidence should be solved by our more general account of what legal proof requires. [...] Whatever explains why evidence in [criminal cases] fails to prove guilt beyond any reasonable doubt, there must be some feature of this evidence that also explains why the evidence in [civil cases] fails to prove liability by a preponderance of the evidence” (Moss 2018: 207)¹²

I believe there are *prima facie* reasons to think that the descriptive, the explanatory, and the normative dimensions of the Epistemic Project are mistaken.

First, epistemologists in this debate often overestimate the extent to which the legal practice treats purely statistical evidence as insufficient for meeting the standard of proof, thereby failing to provide an accurate description of how the law works (Di Bello (2019); Ross (2019); Spottswood forthcoming). Similarly, one

⁹ Spottswood (forthcoming) has also recently drawn these distinctions.

¹⁰ Here I focus on the civil standard of proof “preponderance of evidence” (or “balance of probabilities”). Other standards of proof include: “beyond any reasonable doubt” (that applies to verdicts in criminal cases), “reasonable suspicion” (that applies to police searches). Discussion of the proof paradox for the criminal standard of proof traditionally involve the “Gate-crasher Case” (Cohen 1977, Smith 2020), and the “Prison Yard” (Nesson 1979, Redmayne 2008).

¹¹ Cold-hit DNA evidence refers to when a DNA profile found in the crime scene matches a known profile in a DNA database. This results in an identification of a suspect. Cold-hit DNA evidence is classified as statistical evidence because, roughly put, there’s a possibility of a random match. That is, there could be another person in the population (who might be the actual perpetrator) whose profile would match the DNA sample of the crime scene, but which is not in the database.

¹² In focusing on the criminal case, Littlejohn (2017) is an exception.

might worry that tackling the proof paradox as if it were a *descriptive* project might lead to overestimate the extent to which testimony alone is sufficient for meeting the standard of proof, e.g., in criminal cases.¹³

Second, epistemologists have overestimated the extent to which lottery cases and legal cases have similar structures and generate similar intuitions. Let's assume intuitions to be robust in the legal cases. Let's grant that our reluctance to find liability/guilt on the basis of statistical evidence only, and our willingness to find liability/guilt on the basis of testimony is in fact strong and widely shared. However, things are more complicated when considering the corresponding lottery cases, depending on whether we're interested in ascribing *justification* or *knowledge* to lottery propositions. Although lay people might tend to ascribe justification (and maybe knowledge) in cases like LOTTERY NEWSPAPER, things are less straightforward when it comes to evaluating cases like LOTTERY. For instance, the empirical studies conducted by Ebert, Smith, and Durbach (2017) show that, when evidence is statistical, people tend to attribute more *justification* to lottery propositions when the lottery is large, while the same effect hasn't been found in cases of *knowledge* attribution. Similarly, not only are intuitions not clear and robust in the epistemic case, but there's also *theoretical* disagreement over whether beliefs in lottery propositions deserve any positive epistemic status. While most would say that we cannot *know* propositions like "my lottery ticket is a loser", some are happy to say that we can *justifiably* believe them (Foley, 1987, McGlynn, 2012), while others would deny even justification (Smith 2010). Given intuitions aren't very robust in the epistemic cases, we should be suspicious of a project that heavily relies on what our intuitions are in LOTTERY in order to account for BLUE BUS. Finally, we shouldn't underestimate the extent to which different variations of the Proof Paradox might come with different intuitions, depending on the details of the case and depending on what kind of statistical and non-statistical evidence we are presented with. For it's plausible to think that testimonial evidence might raise some moral and social concerns that evidence provided in the form of CCTV footage doesn't. Similarly, it's also plausible to think that different kinds of *statistical* evidence prompt different kinds of intuitions. For instance, it's not clear that statistical evidence in the form of cold-hit DNA evidence generates the same reluctance as statistical evidence like the ratio of blue to red buses does. If anything, juries might be prone to ignore the possibility of a random match altogether and, when instructed by an expert about it, it wouldn't be surprising if juries would underestimate the extent to which it's a genuine possibility. After all, as Weinberg et al. (2010) have shown, the "Wells' effect" mentioned in §2.1. is less pronounced when the given probabilities are very low.

However, one might insist that, regardless of how the current legal system operates, and regardless of how robust our intuitions are across cases, the *normative* project is nevertheless valuable. Crucially, the

¹³ As Smith points out (2018: footnote 4), in *Scots Law* the "rule of corroboration" prohibits convictions on the basis of a single piece of eyewitness testimony.

above considerations should make us suspicious of the normative dimension of the project as well. In fact, it's plausible that the different strength of the intuitions we have in LOTTERY and BLUE BUS might be the result of a substantial, yet overlooked, asymmetry between epistemic and legal cases more generally. Cases like LOTTERY invite the reader to assess whether a subject's *belief* deserves any epistemic good status. Cases like BLUE BUS invite the reader to assess whether a *verdict* is *just* or *unjust*.¹⁴ But then why should we have one evaluative standard for different objects of evaluation? While this doesn't represent a decisive argument against the Epistemic Project, this asymmetry should make us at least suspicious of the normative dimension of this project as well.¹⁵ The following sections can be seen as corroborating this suspicion. However, before moving on, let me clarify three assumptions I'll be working with. First, in this paper, I assume that Legal Probabilism is wrong. Second, for reasons of simplicity, I focus on the civil version of the proof paradox. Finally, I bracket the issues presented in this section. That is, I assume that BLUE BUS and BLUE BUS TESTIMONY accurately *describe* how statistical and testimonial evidence are treated in the civil litigation; second, I assume they generate strong and shared *intuitions*. Finally, I assume the following *normative* claims:

Insufficiency of Statistical Evidence (INSUFF-S): we should not find the Blue Bus Company liable on the basis of statistical evidence alone in Blue Bus;

Sufficiency of Testimonial Evidence (SUFF-T): we should find the Blue Bus Company liable on the basis of testimony alone in BLUE BUS TESTIMONY.

My argument against the Epistemic Project will thus take its weakest form, insofar as it won't question any of its main assumptions.

With these preliminaries in place, in the next section, I spell out three weak requirements that every successful epistemic account should meet. In §4, I consider three recent epistemic accounts and I argue that, even while granting the foregoing assumptions, they still fail to meet these requirements. In §5, I provide reasons to be pessimistic about a successful epistemic account of the proof paradox.

3. Three Requirements for Success

¹⁴ For recent considerations in a similar spirit see Backes (2020). While Backes questions the “law-to-epistemology direction of exchange” while remaining sympathetic to the “epistemology-to-law” direction, this paper questions the plausibility of the latter.

¹⁵ Epistemologists often try to explain the inadequacy/adequacy of non-individualised/individualised evidence *outside* the legal domain too. Gardiner (2020b) argues that we shouldn't explain legal and non-legal cases in a unified fashion. Although I agree with Gardiner, this paper argues for a broader claim: appealing to these “epistemic qualities” is also unsatisfying when dealing with *legal* cases. In §4.3, I consider and reject Gardiner's epistemic account.

What should we expect from a *successful* epistemic account? The first requirement follows from the aim of the Epistemic Project itself. As mentioned above, in dealing with the Proof Paradox, the Epistemic Project can be seen as offering an alternative to Legal Probabilism, which understands standards of proof probabilistically. Note, however, that Legal Probabilism and the Epistemic Project rest on two fundamentally different methodologies. On the one hand, defenders of Legal Probabilism start from a value taken to be relevant in the legal context (i.e., accuracy), and then derive how the law should operate in a way that promotes this value. It's not surprising that defenders of Legal Probabilism have often denied that the Proof Paradox is a genuine paradox, while insisting that statistical evidence alone *should* be enough for findings of liability (Schoeman 1987; Ross 2019; Hedden and Colyvan, 2019; Papineau 2019: §6). The Epistemic Project, instead, takes some desirable features of the legal practice as a starting point, and then identifies a value (or set of values) that can explain such features. In addressing the Proof Paradox, epistemologists start from the assumption that we should not find the Blue Bus Company liable on the basis of naked statistical evidence but we should find the Blue Bus Company liable on the basis of testimony alone. The question they aim to answer is not *whether* INSUFF-S and SUFF-T are true. Instead, they ask: which (epistemic) values can we appeal to in order to explain *why* these normative claims are true? With this in mind, we can identify a first desideratum that a successful epistemic account should meet. Whichever "epistemic quality" epistemologists decide to use to interpret the standard of proof, attending to such non-probabilistic epistemic quality has to explain both INSUFF-S and SUFF-T. That is, a successful epistemic account is one which identifies an epistemic value V, which is not probability, which statistical evidence lacks and legal testimony has. Call this the **Value Requirement**.

The second and third desiderata stem from considerations about the nature and the role of standards of proof more generally. As explained above, the Epistemic Project offers necessary epistemic conditions for meeting the standard of proof. However, while epistemologists have focused on the parallel between standards for *epistemic justification/knowledge* and standards of *proof*, if we want to offer a satisfying non-probabilistic interpretation of the standards of proof, we should take a step back and ask a more fundamental question: what are standards of proof for? Assuming it's impossible to achieve absolute certainty about a verdict, a traditional way of answering this question takes standards of proof to deliver rules that guide our decision-making in how to manage the expected cost of legal error in a just way (Pardo 2013: 559; Stein 2005: 138, Hamer 2014). This, in turn, is traditionally understood against the expected utility theory framework (Nance 1998: 622). Roughly put, to obtain the expected cost of legal error, we multiply the probability that the finding is erroneous by the cost of it being mistaken. The higher the cost of legal error, the more demanding the standard of proof. This explains why the criminal standard of proof is much higher than the civil standard: a mistaken conviction is regarded as much worse than mistakenly finding someone liable. Considering the details of expected utility theory is beyond the

scope of this paper.¹⁶ What matters for our purposes is that standards of proof should be understood as delivering rules for managing the legal risk of an unjust verdict. But what does it mean to manage risk in a *just* way? Here's an answer that I believe everyone – including those who reject the expected utility framework– should endorse: legal risk should be managed in a way that doesn't clash with the values and functions the legal trial wants to promote. Consider a couple who, in order to minimise the risk that their kids hurt themselves, never let them leave the house. Even if we assume this is an effective way of minimising the risk that they hurt themselves, this way of managing risk would clash with other important values and functions that parents should promote in raising their children, e.g., their freedom and independence. In the long run, this way of managing risk would be unsustainable. Similarly, however epistemologists decide to interpret the standards of proof, these should deliver rules for legal risk-management that are in line with the values and the functions of the legal trial. Alternatively, *if* the target epistemic interpretation of the standards of proof clashes with the values underpinning the legal trial, then the defender of such account has to explain why this revision of the standards is worth having. Consider again the civil standard of proof. Assume that there is an epistemic value V (different from probability) such that attending to V explains both INSUFF-S and SUFF-T. To be a serious and successful competitor of Legal Probabilism, the defender of the epistemic account has to explain why the law should care about such value V, and why INSUFF-S and SUFF-T are worth saving on pain of revisions. In sum, meeting the Value Requirement won't be enough. Not only a successful epistemic revision of the standard of proof explains INSUFF-S and SUFF-T, but it does so in a way that is in line with the function and values of the legal trial. Call this the **Functionalist Requirement**.

However, meeting the **Functionalist Requirement** is also insufficient. Given that standards of proof are expected to *guide* us in deciding whether to find for the defendant or the plaintiff, they should prescribe rules for managing legal risk that are feasible. Call this the **Feasibility Requirement**. Consider the rule “Ride a bike only if you're certain that you won't ever get hit by asteroids”. This rule might promote the value of safely riding a bike, but it cannot be implemented by the average cyclist (assuming that the average cyclist lacks relevant knowledge about asteroids). This rule would make cycling virtually impossible. Similarly, assume for the sake of the example that the central value underpinning the legal trial is the accuracy of the verdict. A standard of proof that prescribes that we should find in favour of the plaintiff only if the probability that the defendant is liable given the evidence is 1 surely serves the value of accuracy. However, it would make finding someone liable very difficult. In this sense, a standard

¹⁶ For a recent objection to the expected utility interpretation of standards of proof see Gardiner (2017)

that fails to meet the Feasibility Requirement will trivially fail to meet the Functionalist Requirement as well.¹⁷

To sum up, in offering an interpretation of the standards of proof (SOP) in terms of a non-probabilistic epistemic value (V), a successful epistemic account will have to meet the following three requirements:

Value Requirement: the interpretation of SOP in terms of V is such that appealing to V explains why INSUFF-S and SUFF-T.^{18 19}

Functionalist Requirement: the interpretation of SOP in terms of V delivers rules for managing legal risk that are in line with the values and the functions underpinning the legal trial.

Feasibility Requirement: the interpretation of SOP in terms of V delivers rule for managing legal risk that can be feasibly implemented.

In what follows, I consider three recent epistemic accounts that explicitly tackle the proof paradox by providing a theory of risk and risk management alternative to the probabilistic one underpinning Legal Probabilism. Despite their differences, these epistemic accounts all share the idea that managing legal risk is not a matter of minimising error. Instead, managing legal risk requires ruling out modal risk (Pritchard 2018), normic risk (Ebert et al. 2019; Smith 2018), or all the relevant alternatives to the litigated claim (Gardiner 2019 2020). Crucially, I argue that these accounts struggle to meet the three above-mentioned requirements. In §5, I will provide reasons to be pessimistic about the prospects of finding a successful epistemic account.

4. The Modal, the Normic, and the Relevant

4.1. Modal Risk

¹⁷ For example, Littlejohn (2017) and Duff (2012) have argued that meeting the criminal standard requires knowledge that the defendant is guilty. One might worry that these standards are too idealised to be realistically achieved. See, for instance, Blome-Tillmann (2017).

¹⁸ As I explain in §5, I believe that in fact we should *not* try to explain INSUFF-S and SUFF-T by appealing to an epistemic value. However, this is exactly what the Epistemic Project primarily aims to do. My argument against the Epistemic Project takes its weakest form: *if* the Epistemic Project is successful, then it meets the Value Requirement. And yet, as the next section shows, it's not clear that the Epistemic Project can meet this weak desideratum.

¹⁹ Three things worth clarifying here. First, note that the Epistemic Project wants to identify a *necessary* epistemic condition other than probability for meeting the standard of proof. This means that other epistemic and/or non-epistemic conditions might still be required for meeting the standard of proof. However, what matters here is that, according to this project, what explains the different treatment between testimony and statistical evidence in cases like Blue Bus/Blue Bus Testimony can be fully explained by appealing to some *epistemic* value (as opposed to moral, practical, or social). Second, even though, as far as I know, epistemologists explaining the proof paradox have provided *one* epistemic value to account for the different treatment of statistical evidence and testimony in proof paradox-cases, I believe it is consistent with the spirit of the Epistemic Project to take V to stand for a combination of epistemic values. Thanks to an anonymous referee for pushing me to clarify this issue here.

The first epistemic interpretation of the standards of proof I'll consider is the one put forward by Pritchard. According to Pritchard, the notion of risk that is relevant to both epistemology and the law is not a probabilistic one (Pritchard 2018 2016 2015). Instead, on his modal account of risk, the risk of a (negative) event is determined by its *modal closeness*. That is, it depends on whether it could have easily occurred. This is supposed to explain why you don't know that your lottery ticket is a loser on the basis of the odds involved. Knowledge requires ruling out modal risk, and yet probability and modal closeness can come apart. After all, even if it's very probable that your ticket is a loser, the winner could easily be you! But how should we understand "easily"? Although Pritchard believes it's possible for the probability of a proposition and its modal profile to come apart, he seems to nevertheless take modal closeness to be an *objective* standard, one related to "how much things need to change in the actual world in order for that event to occur". As he puts it:

"The point is that we naturally order possible worlds, and thus the possible events that obtain in those worlds, in terms of their similarity to the actual world, where similarity is determined by how much needs to change in the actual world in order to get to this possible world where the target event occurs". (Pritchard 2018: 112)

How can appealing to "objective modal risk" help us explaining the Proof Paradox? Roughly put, meeting the standard of proof requires the available evidence to rule out modal risk of mistakenly finding the defendant liable. Regardless of what the probabilities involved are, the standard of proof is met if and only if it could not be an *easy* possibility that the defendant is wrongfully found liable. This (allegedly) predicts that, given the testimony that a blue bus hit Mr. Brown, there's a very low modal risk that the Red Bus Company is actually the one responsible for the accident. This is because, for the testimony to be false, it would require a lot of things to change in the actual world, such as, "the Red Bus Company involving one of their buses being painted blue in order to create problems for the rival company" (Pritchard 2018: 117-118).

For the sake of this paper, I will grant Pritchard that objective modal closeness and probabilities can come apart.²⁰ Crucially, by defining standards of proof in terms of objective modal risk, Pritchard's account will prescribe rules for risk-management that are not feasible. This is because we often don't have access to the objective features of the actual world, thereby making it difficult to assess the degree of objective similarity between the actual world and a possible world. To see more clearly why this is problematic in the legal context, consider the following case:

HALLY

Mr. Brown is run over by a bus on Montgomery Street. He couldn't see which bus hit him. Hally, a bystander, testifies that she saw a blue bus hitting Mr. Brown. The expected

²⁰ See Yang (2019) for a dissenting view.

reliability of eyewitness testimony is approximately 70%. The only available evidence is Hally's eyewitness testimony. There's no evidence against her reliability. However, unbeknownst to the court and unbeknownst to Hally as well, she's prone to colour hallucination.

Given it is part of Hally's cognitive architecture that she's prone to colour hallucination, the world in which Hally is mistaken is modally very close to the actual world. If Pritchard's account were correct, the Blue Bus Company should *not* be found liable in this case. But this doesn't seem plausible. And just as we don't have access to the objective features of the actual world, we can't access the features of close possible worlds either. And yet, as recently pointed out by Ebert et al (2019), Pritchard's account predicts that determining whether a negative event (finding an innocent liable) is at high risk of occurring will require to know in advance whether the negative event occurs in a close or far possible world. In other words, given our inaccessibility to the objective ordering of possible worlds, Pritchard's modal account clashes with the guiding role we expect standards of proof to play. It delivers rules for risk management that don't meet the Feasibility Requirement.

Perhaps there's a different way of cashing out the notion of similarity, one that is more in line with our intuitions of what counts as "similar" and "risky". Call this the "subjective modal risk" account. Pritchard himself seems to suggest something similar. In fact, he stresses how his modal account of risk is rooted in our everyday assessment of events, and is in line with folk intuitions and judgments as to what counts as a close and risky event. He explicitly takes this to be a virtue of the modal conception of risk over the probabilistic one, which he takes to be highly theoretically driven (Pritchard 2018: 113). To show that this is the case, he invites the reader to consider a paradigmatic case in which our everyday risk assessment largely diverges from the actual probability that the negative event would happen: whether it's riskier to drive a car rather than to take a train. In this case, driving a car is judged to be less risky than taking the train, despite involving a higher probability of danger. Appealing to "subjective modal risk" makes sense of these intuitions.

How is "subjective modal risk" supposed to have an advantage over "objective modal risk" when it comes to the law? First, if the notion of modal closeness (and thus risk) is rooted in our intuitions, then the corresponding ordering of possible worlds will be easily available to us. Understanding risk in terms of subjective modal closeness delivers rules for risk-management that seem to easily meet the Feasibility Requirement. Second, in order to account for the intuitions underpinning BLUE BUS TESTIMONY and HALLY, Pritchard could say that, regardless of what the *objective* features of the actual and possible worlds are, cases in which the testifier is mistaken are usually *perceived* as far-fetched and distant possibilities: testimony does rule out subjective modal risk. The drawback, however, is that this notion of modal risk is too subjective to usefully inform the notion of *legal* risk and the legal standards of proof

more generally. Note that, while explicitly making the *normative* claim that we *should* understand standards of proof in terms of modal legal risk, Pritchard is aware of how “various cognitive biases have a role to play in leading subjects to make these assessments of [modal] risk” (2018: 113).²¹ This, however, raises a number of important methodological issues. Consider the following case:

MARIO

Mr. Brown is run over by a bus on Montgomery Street. He couldn't see which bus hit him. Mario, a bystander, testifies that he saw a blue bus hitting Mr. Brown. The expected reliability of eyewitness testimony is approximately 70%. The only available evidence is Mario's eyewitness testimony and there's no evidence against Mario's reliability or sincerity. However, the judge has recently seen a movie where the villain was played by an Italian actor, and she can't help but thinking that Mario could also easily be a vicious person who's committing perjury.

First, note that it's plausible to take BLUE BUS TESTIMONY as being relevantly similar to MARIO. MARIO's testimony is the only testimony available, he has passed cross-examination, and we have no reason to doubt his credibility. It's plausible to say that, if we should not find the Blue Bus Company liable in MARIO, then we should find for the defendant in BLUE BUS TESTIMONY as well. And vice versa, if we think we should find the Blue Bus Company liable in BLUE BUS TESTIMONY, we should say the same about MARIO. Denying equal treatment of MARIO and BLUE BUS TESTIMONY means that how we should manage legal risk can depend on what movie the judge has watched the night before. This unfair treatment of the testifier would clash with the functions and values underpinning the legal trial.

With this in mind, what should we say about MARIO? If the standard of proof is met when the evidence rules out subjective modal risk, then Pritchard's account predicts that the judge should disregard MARIO's testimony and find in favour of the defendant. The availability bias, for instance, perfectly explains why the world in which MARIO is mistaken would be perceived as a *close* (and thus “risky”) possible world. But this seems wrong. After all, we surely wouldn't want our cognitive biases to inform the standards of proof, thereby legitimising morally problematic verdicts (racist, sexist, and classist verdicts).²² The problem with Pritchard's methodology is that he uses the modal account of risk both to explain our intuitive risk-assessment and to inform the notion of legal risk that should be used to define standards of proof, while aware that our intuitive risk judgments are driven by our cognitive biases.

²¹ The *loci classici* on risk perception are Kahneman & Varey (1990) and Kahneman (2012).

²² One might say that this is not a problem for Pritchard's account *per se*, but it's a problem for any account that is applied in a morally problematic way, for any account that is applied in an ethnically biased way will give rise to cases of injustice. However, note that the problem with Pritchard's account is not that it is *applied* in a morally problematic way. Rather, it is the very notion of modal closeness that underpins his account that is hostage to morally problematic biases. Thanks to an anonymous reviewer for raising this issue.

Cashing out standards of proof in terms of subjective modal risk delivers rules for legal risk-management that clash with the Functionalist Requirement.

However, if we think that the judge *should* find the defendant liable on the basis of MARIO's testimony, then appealing to subjective modal closeness won't explain why this is the case given that MARIO's testimony doesn't rule out subjective modal risk that the defendant is innocent. Unless we want to deny equal treatment of MARIO and BLUE BUS TESTIMONY, the same will apply, *mutatis mutandis*, to BLUE BUS TESTIMONY. Pritchard's interpretation of the standards of proof in terms of subjective modal risk would fail to meet the Value Requirement.

Let's take stock. There are two ways in which we can understand the notion of similarity underpinning Pritchard's modal account of legal risk and the corresponding interpretation of standards of proof. Either we take similarity to be an objective relation between the actual world and possible worlds. Or we understand it as a function of the heuristics informing our intuitive risk judgments. Both accounts are problematic. The former prescribes rules for managing risk that fail to meet the Feasibility Requirement. The latter faces a trilemma: i) either it predicts that we should manage legal risk in the same way in both MARIO and BLUE BUS TESTIMONY and that the standard of proof (SOP) is met in both cases, but then appealing to subjective modal closeness would fail to meet the Value Requirement; ii) or it predicts that, although MARIO and BLUE BUS TESTIMONY deserve equal treatment, SOP is *not* met in either case, but then the subjective modal closeness account would trivially fail to meet the Functionalist Requirement;²³ iii) alternatively, Pritchard could say that legal risk should be managed differently in BLUE BUS TESTIMONY and MARIO: SOP is not met in MARIO, but it's met in BLUE BUS TESTIMONY. However, as we have seen, treating these two cases differently would also fail to meet the Functionalist Requirement. Attending to modal considerations won't get us to the heart of the problem. We need to look elsewhere.

4.2. Normic Risk

Pritchard's modal account of risk orders possible worlds in terms of how *similar* they are/are perceived to be to the actual world. However, possible worlds can also be ordered in terms of how *normal* they are with respect to the actual world. This idea underpins the "normic" account of risk (Ebert et al. (2018)). Roughly put, an event described by a proposition *p* is at risk of happening relative to a subject's evidence

²³ This is because, given we are assuming that INSUFF-S and SUFF-T are two desirable features of the legal practice, any epistemic account that entails the rejection of INSUFF-S and SUFF-T will trivially clash with the legal practice.

E if and only if the case in which [E and p] is more normal than the case in which [E and not-p], where “more normal” means that the “circumstance in which E is true and P is false requires more explanation than the circumstance in which E and P are both true.” (Smith 2016: 40; 2018: 16).²⁴

On this view, managing legal risk well doesn’t involve minimising error. Instead, it means that the available evidence has to rule out *normic* risk of mistakenly finding the defendant liable. That is, legal risk is managed appropriately (and the standard of proof is met) only if it’s more normal, given the evidence, that the defendant is liable rather than innocent. That is, only if the case in which the defendant is innocent would be abnormal given the evidence and it would thus require some extra explanation.

But what kind of *extra* explanation would be required in *abnormal* cases? Here’s what Smith says:

“[A case is abnormal when,] if one’s belief turns out to be false, then the error has to be *explicable* in terms of disobliging environmental conditions, deceit, cognitive or perceptual malfunction, etc. In short the error must be attributable to *mitigating circumstances* of some kind and thus *excusable*, after a fashion.” (Smith 2016: 41)

Sometimes, when we describe an event as “normal”, we mean that it’s *frequent*. However, it’s important to stress that, the normic account can have an advantage over the competition (i.e., the probabilistic account) only insofar as “normal” is *not* meant to be a claim about frequency (Ebert et al. 2018: 12). Consider LOTTERY. On this account, the reason why I lack justification to believe that my lottery ticket is a loser merely on the basis of the odds involved is that both the world in which I win the lottery and the world in which I lose it can be exhaustively explained by appealing to the random nature of the event (Smith 2010; 2016: Ch. 3). By contrast, I have justification to believe that my ticket is the loser on the basis of the newspaper report because the world in which I win the lottery ticket given the newspaper’s report (allegedly) cannot be exhaustively explained by appealing to considerations about probability. Putting this notion at work in the legal context, we have that the available evidence [E] rules out normic legal risk that the defendant is innocent [q] if and only if the case in which [E and not-q] cannot be exhaustively explained by appealing to considerations about the probability that not-q.

Is this account better than the modal account? By relativizing the ordering of normal worlds to the subject’s evidence, the normic account delivers rules for risk-management that, unlike Pritchard’s objective modal account, seem to better accommodate the Feasibility Requirement. Crucially, it’s not clear whether it meets the Value Requirement, which, remember, requires explaining both INSUFF-S and SUFF-T. Here’s what Smith (2018) says: in BLUE BUS, the statistical evidence that 90% of the buses belong to the Blue Bus Company does not rule out the normic legal risk that a Blue Bus Company is innocent. For if it turned out that a red bus hit Mr. Brown, we could explain this event merely by appealing

²⁴ This notion of normic risk stems from Smith’s notion of *normic* (as opposed to probabilistic) evidential support (2016).

to statistical facts. By contrast, in BLUE BUS TESTIMONY, the testimony (allegedly) rules out the normic legal risk that the blue bus is innocent: if it turned out that a red bus hit Mr. Brown, we wouldn't be able to explain the falsity of the testimony merely by appealing to statistical facts. Instead, we would need to appeal to some extra mitigating circumstances. To put it in Smith's words:

“[I]f it turned out that the bus involved was not owned by the Blue Bus Company, in spite of the eyewitness testimony, then there would have to be some accompanying explanation – the eyewitness was hallucinating or lying or had a fabricated memory” (Smith, 2016: 39).

The problem with appealing to normic risk is that it's difficult to think of BLUE BUS TESTIMONY as a case in which appealing to statistical considerations cannot explain the target error-possibility. In other words, if Pritchard's subjective modal account is too psychologised, the normic account, as it is, seems to be too closely related to considerations about the *probability* of the negative event occurring. Consider BLUE BUS TESTIMONY again. The possibility that “the eyewitness was hallucinating or lying or had a fabricated memory” are exactly those same facts that are considered when evaluating the overall *reliability* of eyewitness testimony. If this is so, then, if the testimony turned out to be false, it's not clear why any extra explanation would be required, besides considerations that have to do with how *frequent* events like false testimonies are.

There's another related problem. Consider a proposition *p*. As Backes (2018) has recently pointed out, as soon as one's body of evidence *E* contains an explanation of why *p* may turn out to be false, then the case in which [*E* and not-*p*] is not abnormal anymore. After all, the case in which *p* is false would no longer call for an extra explanation. If anything, the case in which [*E* and not-*p*] would make perfect sense, given that the possible explanation of why *p* could be false is available. Why is this relevant for our purposes? Because the possible reasons why and when eyewitness testimonies may be false are widely known and available, as the Innocence Project²⁵ and the influential studies conducted by Loftus (1996) show.

Where does this leave us? Smith's interpretation of standards of proof in terms of how they rule out normic legal risk delivers rules for managing legal risk that, although feasible, they fail to meet the Value Requirement. While it accounts for INSUFF-S, it struggles to explain SUFF-T: testimony doesn't rule out the normic legal risk that the Blue Bus Company is innocent.

4.3. Relevant Alternatives

²⁵ <https://innocenceproject.org/>

Finally, let's consider Gardiner's "relevant alternative" account of risk (Gardiner 2019 2020). Here, I will argue that, at least when used to interpret standards of proof, Gardiner's relevant alternative framework delivers rules for managing legal risk that, although they can be feasibly implemented, they fail to meet either the Value or the Functionalist Requirement.

Drawing on Lewis' relevant alternative account of *knowledge*, on which, knowledge that p requires ruling out relevant alternatives to p (Lewis 1996), Gardiner has recently argued that managing the risk of p being false requires ruling out relevant alternatives to p (Gardiner 2019 2020). Standards of proof should thus not be understood as corresponding to different probability thresholds to which the claim needs to be established. Instead, standards of proof should be understood in terms of the relevant alternatives that cannot be properly ignored. The more demanding a standard of proof is, the more alternatives to the litigated claim will be considered relevant and, therefore, can't be ignored. As Gardiner says:

"Claim p is established to a legal standard L only if the evidence adduced rules out the L-relevant [alternatives]" (2019: 300)

Before showing why Gardiner's proposal is problematic, some clarifications on the terminology is needed.

First, what is an *alternative*? An *alternative* to p is a proposition incompatible with p. Take p to be the proposition that the animal in the pen is a zebra. Alternatives to p include, for instance, that the animal is a bird, or that it's a mule disguised as a zebra. Take q to be the proposition that the animal in the pen *looks like* a zebra. The error-possibility that the animal is actually a mule disguised as a zebra is *not* an alternative to q: q is perfectly compatible with the fact that it's a mule disguised as a zebra. For each proposition, there can be infinitely many alternatives and each alternative can be further divided into sub-alternatives. But surely we can't rule out *all* possible alternatives. So here's where the *relevant* component comes in.

A *relevant* alternative to p is an alternative that *has* to be ruled out, namely, that cannot be properly ignored (2019: 294). But what makes an alternative *relevant*? Unfortunately, Gardiner doesn't offer us a clear set of desiderata. Instead, she draws on intuitions: "Some error possibilities are farfetched, they can be properly ignored. Others cannot be properly ignored - they seem important, relevant, reasonable - they must be ruled out" (2019). In fact, she explicitly offers the relevant alternative account of risk as a "skeletal

structure that can be combined with various accounts of what determines remoteness of possibilities. These include whether the alternative is true, normal, or statistically probable” (2020: 11-12).²⁶

Finally, when does evidence *rule out* a relevant alternative? According to Gardiner, ruling out an alternative requires that the evidence *addresses* the relevant incompatible alternative. But what does it mean for the evidence to *address* a relevant alternative? To put it in Gardiner’s words:

“[A]n error possibility is “addressed” by the evidence when the possibility is rendered into branched alternatives and *each* node is either incompatible within the evidence or lies beyond the threshold of what qualifies as relevant.” (2019: 295). (Italic is mine)

In a nutshell, evidence addresses a *relevant* error possibility when it’s incompatible with *all relevant* alternatives and sub-alternatives to p.²⁷ Note that, on Gardiner’s view, ruling out alternatives doesn’t require us to *think* about whether our evidence is incompatible with some relevant error-possibilities. Rather, ruling out alternatives is something that happens spontaneously and effortlessly. Evidence addresses alternatives automatically all the time (2020: §3).

By relativizing risk to one’s evidence, and by requiring that the evidence rules out only the *relevant* alternatives, this account seems to deliver rules of risk management that can be implemented, thereby meeting the Feasibility Requirement. However, for this account to be satisfying, the fact that testimony rules out all relevant alternatives should explain both why it’s not permissible to find the Blue Bus Company liable on the basis of statistical evidence alone (INSUFF-S), and why we should find the Blue Bus Company liable on the basis of testimony alone (SUFF-T). Can this framework do that? First note that, by remaining silent on what makes an alternative *relevant*, Gardiner’s account risks losing motivation. In particular, if her account allows us to define relevance in terms of what’s “statistically probable”, it’s unclear what advantage Gardiner’s epistemic account has over Legal Probabilism. If anything, understanding relevance in terms of probability would generate the paradox, just as Legal Probabilism does. But, assuming that an informative non-probabilistic conception of “relevant” exists, here’s how Gardiner explains the inadequacy of statistical evidence: the statistical evidence that 90% of the buses in town belong to the Blue Bus Company is compatible with the relevant alternative that a red bus hit Mr Brown. In other words, the available statistical evidence fails to address the relevant alternative (2019: 315). But what about the *adequacy* of testimony? Does the testimony in BLUE BUS TESTIMONY address all the relevant alternatives? I believe it does not. The problem lies in the way in which Gardiner

²⁶ Note that, on Gardiner’s view, and unlike Lewis, merely mentioning or considering an alternative doesn’t make it relevant (2019: 7; 2020: 12).

²⁷ Thanks to Lilith Newton for stressing this point to me.

understands the notion of “ruling out”. For, on her view, whether the evidence addresses the relevant alternative merely depends on whether the evidence is *incompatible* with the relevant alternative. Crucially, there are going to be many (plausibly) relevant alternatives that are compatible with the fact that the witness says that she saw a Blue Bus hitting Mr Brown: the eyewitness might be lying or misremembering. As it stands, this account doesn’t meet the Value Requirement.

Perhaps one could further lower the civil standard of proof by restricting the sphere of relevant alternatives. However, as Gardiner herself points out, “[the error possibilities] that are sufficiently typical must be taken seriously by the court” (2019: 304). If the possibility that the eyewitness misremembers or is lying doesn’t count as a *typical* error possibility, then it’s not clear what does.

Another option would be to say that testimony fails to address the relevant alternative only insofar as we are thinking of testimonial evidence as the proposition that [the testifier said that the blue bus hit Mr Brown]. Instead, Gardiner might argue, we should think of it as the proposition that [the blue bus hit Mr Brown]. Once we take our evidence to be that [p] (rather than the proposition that [she says that p]), then, in order for our evidence to rule out legal risk, [p] has to address the relevant error possibilities. And yet, Gardiner could say, our evidence [p] is in fact *incompatible* with the error possibility that she’s misremembering. So, contrary to first appearances, we have the result that testimony rules out legal risk that a red bus hit Mr Brown.²⁸

This strategy would make sure the Value Requirement is met, but it would fail to meet the Functionalist Requirement. For even if we grant that this is how we should think of testimony in *ordinary* circumstances, this view clashes with how evidence is (and should be) treated and assessed in the legal context. Treating testimonial evidence as the claim that p (as opposed to the claim that she said that p) fails to explain the defendant’s right to cross-examine the witness in order to question or cast doubts on the witnesses’ credibility (Dennis 20017, §14, C). This practice is difficult to explain unless we assume that, in legal contexts, the error possibilities related to the witness’ reliability and sincerity are taken to be relevant.

²⁸ This option is inspired by Gardiner’s strategy for explaining why we should believe victims of sexual assault (2020: §9). However, the story she offers there is slightly different. She argues that, in cases of rape accusations, people mistakenly consider the claim “she’s telling the truth about being raped” rather than the claim “she was raped”. By doing so, people mistakenly consider error possibilities, e.g., “she’s lying”, as *relevant* alternatives to the target claim thereby failing to rule out such alternatives. Instead, Gardiner argues, once we consider the claim “she was raped”, the alternative that the testifier is lying is typically remote and can be ignored. Two points here. First (and this is supportive of her strategy), it’s not clear why Gardiner appeals to the fact that the testimony is compatible with the *irrelevant* alternative “the testifier is lying”. For her view predicts a more straightforward explanation: once the target evidence people acquire is “she was raped”, the hearer’s evidence (which now includes the proposition “she was raped”) is *incompatible* with the error possibility that she’s lying. However – and this is the second critical point – note that, unless we have a better characterisation of what counts as a *relevant* alternative, appealing to probabilities here would perfectly explain why we should believe the victim. (Simion 2020: <https://blogs.cardiff.ac.uk/openfordebate/2020/06/29/youshouldbelieveher/>)

Moreover, this is exactly what we should expect: after all, “p” is exactly the claim that needs to be established to a certain standard *given* the evidence.

In sum, the relevant alternative framework delivers an account of legal risk-management that, either it fails to meet the Value Requirement: the account fails to explain SUFF-T, given that testimony *does not* address the relevant alternatives to the litigated claim; or, upon revising the notion of testimonial evidence, it meets the Value Requirement, but it fails to meet the Functionalist Requirement.

5. Reasons for pessimism and the “value-turn”

At this point, one might wonder whether there’s a better epistemic account available. In the remainder of the paper, I provide reasons to be pessimistic about such prospects.

First, note that, like Pritchard’s objective modal account, many other epistemic accounts will also prescribe rules for risk-management that cannot be feasibly implemented. Consider the “Doxastic Approach” to the proof paradox. On this approach, standards of proof require evidence that generates being in a specific doxastic state towards the proposition that the defendant is liable or guilty (e.g., full belief, knowledge, probabilistic knowledge). But one might worry that, given the non-luminous nature of these mental states, i.e., the fact that we often lack unproblematic access to them (Williamson 2000 ch4; Srinivasan 2015), it would be hard to determine when the standard of proof is met. This clashes with the guiding role we expect standards of proof to play.²⁹

Second, in the previous section, we have seen that appealing to whether the evidence rules out normic risk, modal risk, or relevant alternatives doesn’t explain both INSUFF-S and SUFF-T. At closer inspection, *both* statistical evidence and testimonial evidence lack this quality, and yet, we should nevertheless treat these kinds of evidence differently. My argument against the normic, modal, and the relevant alternative account can be seen as the counterpart of the argument that epistemologists have raised against Legal Probabilism: statistical evidence makes it highly probable that the defendant is liable, and yet it should not be enough for findings of liability. Similarly, I’ve argued that testimony doesn’t rule out normic risk, modal risk, or relevant alternatives, and yet it should nevertheless be enough for findings of liability. In other words, these accounts fail to meet the Value Requirement. Crucially, I believe that other epistemic accounts will face the same problem. Why is that? I believe this is because the Epistemic Project rests on problematic methodology.

²⁹ Enoch et al’s (2012) sensitivity account and Thompson’s (1986) causal account face a similar problem (see Smith 2018: §2).

As explained above, when comparing cases of pure statistical evidence with cases involving testimonial evidence, epistemologists have tried to provide an epistemic value that can explain both the epistemic inadequacy of statistical evidence and the epistemic adequacy of testimony. By doing so they have treated the following questions in a unified fashion:

Why should we not find liability on the basis of statistical evidence alone in BLUE BUS?

Why should we find liability on the basis of testimony alone in BLUE BUS TESTIMONY?

According to the Epistemic Project, whatever answers the first question will answer the second one as well. In other words, the Epistemic Project rests on the idea that what makes a difference in how the law should treat statistical evidence as opposed to testimony, at least in cases like BLUE BUS and BLUE BUS TESTIMONY, rests on the fact that, *ceteris paribus*, testimony has the target non-probabilistic epistemic quality while statistical evidence doesn't. Crucially, the failure to separate these two questions has led epistemologists to overlook the complex yet essential legal and social dimension that arises once we introduce testimonial evidence (as opposed to statistical evidence), and that constitutes the framework within which these puzzles arise in the first place. For given the intrinsically social dimension of testimonial evidence, we should expect the *justness* of verdicts based on testimony to be at least partly sensitive to its social nature. That is, we should expect some non-epistemic values to be relevant in explaining why, in absence of evidence that speaks against the reliability of the testifier, testimonial evidence can be enough for findings of liability. As long as the Epistemic Project presumes an alienated conception of testimony as a mere source of information (just as any other kind of evidence), I believe we should expect that appealing to other epistemic values will also fail to meet the Value Requirement.

Similarly, by drawing on lottery paradoxes, the Epistemic Project rests on the assumption that a purely epistemic explanation of the proof paradox is possible. However, as anticipated in §2, when dealing with lottery paradoxes, epistemologists are concerned with whether a *belief* is *rational* or *justified* or (if true) knowledge. Proof paradoxes, instead, are concerned with determining when a *verdict* is *just* or *unjust*. While overlooking this asymmetry, epistemologists have taken just or unjust verdicts to be primarily a function of an epistemic quality the evidence has or lacks. In other words, they have assumed that managing legal risk in a *just* way is going to primarily depend on whether the admissible evidence serves some epistemic value, e.g., knowledge, sensitivity, normic support, modal stability. But once we take seriously the idea that standards of proof should deliver rules for managing risk that promote or are at least in line with the functions and values of the legal trial, then, cashing out standards of proof primarily in terms of *epistemic*

values is unlikely to get us very far.³⁰ After all, the civil trial serves and promotes many functions and values that are not purely epistemic in nature, including speed, cost, participation, simplicity, fairness, privacy, accessibility, finality (See Michalski 2018). When understood correctly, standards of proof should deliver rules that allow us to manage legal risk in a way that promotes the various functions and values of the legal trial. Furthermore, once we take seriously the differences between the functions and values underpinning the civil trial and those underpinning the criminal trial, we should be suspicious of any attempt to explain both the criminal and the civil version of the proof paradox in a unified fashion. In other words: cashing out standards of proofs in terms of how they primarily serve an *epistemic* value fails to capture the social, moral, and practical dimension of the law. This doesn't necessarily mean that epistemic values including and other than accuracy, e.g. modal stability and normic support, should be completely irrelevant in how we manage legal risk. The point is rather that we should expect other non-epistemic values to be relevant in determining when a verdict is just or unjust in a way that we shouldn't expect them to be relevant when formulating standards for justification or rationality, and when solving epistemic paradoxes like the lottery paradox. Approaching the proof paradox from a purely epistemic perspective will fail to meet the Functionalist Requirement. Looking at the various epistemic and non-epistemic values the trial wants to promote might be a more promising way to go.

6. Conclusion

The aim of this paper was *not* to solve the proof paradox against the Epistemic Project. That is, this paper did not provide an explanation of the paradox along non-epistemic lines. This would require me to identify *which* non-epistemic values can normatively explain the adequacy of testimony and the inadequacy of statistical evidence.³¹ The aim of this paper was also *not* to provide a knockdown objection to the Epistemic Project. Rather, in a context in which the epistemic literature on the proof paradox is rapidly growing, this paper aimed to do three things: i) by taking a step back from the literature, it identified three weak requirements that any successful epistemic account of the proof paradox should meet; ii) it argued that some of recent and influential epistemic accounts fail to meet these requirements; iii) it provided reasons to be pessimistic about the prospects of finding a more successful epistemic account, while suggesting that the debate on proof paradoxes would probably benefit from undergoing a “value-turn”.

³⁰ Note that this objection might also apply to Legal Probabilism. Here I've been assuming Legal Probabilism to be wrong. However, I further explore how/whether this threatens Legal Probabilism in an unpublished manuscript.

³¹ Enoch et al. (2012) and, more recently, Dahlman (2020) offer incentive-based solutions. I provide a different “institutional” explanation in MS. For examples of moral explanation of the paradox see Wasserman 1991 and Duff 2008. For an overview of other non-epistemic explanations see Ross (2020: Section 4) and Gardiner (2019b).

References.

- Achinstein, P. (2003). *The book of evidence*. Oxford University Press.
- Backes, M. (2018). Normalcy, Justification, and the Easy Defeat Problem, *Philosophical Studies*
- Backes, M. (2020). Epistemology and the Law: why there is no epistemic mileage in legal cases, *Philosophical Studies*
- Blome-Tillman, M. (2017). More likely than not. In A. Carter, E. C. Gordon, & Jarvis, *Knowledge first: approaches in epistemology and mind* (pp. 279-291). Oxford University Press.
- Buchak, L. (2014). Belief, credence and norms. *Philosophical studies*.
- Cohen, J. (1977). *The Probable and the Provable*. Oxford University Press.
- Dahlman, C. (2020). Naked statistical evidence and incentives for lawful conduct. *International Journal of Evidence and Proof*.
- Di Bello, M. (2019). Prood Paradoxes and Normic Support: socialising or relativising?. *Mind*.
- Duff, A. (2008). Dangerousness and Citizenship: Essays in Honour of Andrew von Hirsch. In A. Ashworth, & M. Wasik (Eds.), *Fundamentals of Sentencing Theory: Essays in Honour of Andrew von Hirsch* (pp. 141-163). Oxford University Press.
- Duff, R. A. (2012). 'Presuming Innocence'. In: Principles and Values in Criminal Law and Criminal Justice: Essays in Honour of Andrew Ashworth (L. Zedner & J. V. Roberts eds.). Oxford University Press, Oxford.
- Ebert, P., Smith, M., & Durbach, I. (2017). Lottery judgments: a philosophy and experimental study. *Philosophical Psychology*.
- Ebert, P., Smith, M., & Durbach, I. (2019). Varieties of Risk. *Philosophy and Phenomenological Research*.
- Enoch, D., Spectre, L., & Fisher, T. (2012). Statistical evidence, sensitivity, and the legal value of knowledge. *Philosophy and public affairs*.
- Foley, R. (1987). *The theory of epistemic rationality*. Cambridge, MA: Harvard University Press.
- Gardiner, G. (2017). In defence of reasonable doubt. *Journal of Applied Philosophy*.
- Gardiner, G. (2019). The reasonable and the relevant. *Philosophy and public affairs*.
- Gardiner, G. (2019b). Legal Epistemology. *Oxford Bibliographies Online*.
- Gardiner, G. (2020). Relevance and risk: how the relevant alternative framework models the epistemology of risk. *Synthese*.

- Gardiner, G. (2020b). Profiling and Proof: are statistics safe? *Philosophy*.
- Hamer, D. (2014). Presumptions, standards and burdens: managing cost of error, *Law Probability and Risk*, 13, 221-242
- Haack, S. (2014). *Evidence matters*. Cambridge University Press.
- Hawthorne, J. (2004). *Knowledge and lotteries*. Oxford: Clarendon Press.
- Hedden, B. and Colyvan, M. (2019). Legal Probabilism: a Qualified Defence, in *Journal of Political Philosophy* 27 (4):448-468.
- Kahneman, D. (2012). *Thinking fast and slow*. Penguin.
- Kahneman, D., & Varey, C. (1990). Propensities and counterfactuals: the loser who almost won. *Journal of personality and social psychology*.
- Kyburg, H. (1961). *Probability and the logic of rational belief*. Wesleyan University Press.
- Laudan, L. (2006). *Truth, Error, and Criminal Law: An Essay in Legal Epistemology*, Cambridge University Press
- Lewis, D. (1996). Elusive Knowledge. *Australasian Journal of Philosophy*
- Littlejohn, C. (2017). Truth, knowledge, and the standard of proof in criminal law. *Synthese*.
- Loftus, E. F. (1996). *EyeWitness Testimony*, Harvard University Press.
- McGlynn, A. (2012). Justification as would be knowledge. *Episteme*.
- Michalski, R. (2018). The clash of procedural values. *Lewis and Clark Law Review*, Vol. 22, No. 1
- Moss, S. (2018). *Probable knowledge*. Oxford: Oxford University Press.
- Nelkin, D. (2000). The lottery paradox, knowledge, and rationality, *The Philosophical Review*, V 109, 3.
- Nance, A. (1998). Evidential Completeness and the Burden of Proof, *Hastings Law Journal*
- Nesson, Charles (1979) 'Reasonable Doubt and Permissive Inferences: The Value of Complexity' *Harvard Law Review* 92(6): 1187–1225.
- Papineau, D. (2019). The disvalue of knowledge, in *Synthese*
- Pardo, M. (2013). The nature and law of evidence theory, in *Vanderbilt Law Review*, V 66, I 2, A3.
- Pritchard, D. (2015). Risk. *Metaphilosophy*.
- Pritchard, D. (2016). Epistemic Risk. *Journal of Philosophy*.
- Pritchard, D. (2018). Legal Risk, Legal Evidence, and the Arithmetic of criminal justice. *Jurisprudence*.

- Redmayne, M. (2008). Exploring the proof paradox. *Legal Theory*, 281-309.
- Ross, L. (2020). Recent work on the Proof Paradox. *Philosophy Compass*
- Ross, L. (2019). Rehabilitating statistical evidence. *Philosophy and Phenomenological Research* .
- Roth, A. (2010). Evidence in numbers: when DNA alone is enough to convict. *New yourk university law review*.
- Schoeman, F. (1986). Statistical vs. Direct Evidence, in *Nous*, Vol. 21, No. 2, pp. 178-198.
- Simion, M. (2020). #YouShouldBelieveHer in <https://blogs.cardiff.ac.uk/openfordebate/2020/06/29/youshouldbelieveher/>
- Smith, M. (2010). What else justification could be? *Nous*.
- Smith, M. (2016). *Between probability and certainty*. Oxford University Press.
- Smith, M. (2018). When does evidence suffice for conviction? *Mind*.
- Smith, M. (forthcoming). Against legal probabilism. In R. a. Hoskins, *Truth and Trial*. Routledge.
- Spottswood, M. forthcoming. Burdens of Proof, in, eds. Dahlman, Stein, and Tuzet, *Philosophical Foundations of Evidence Law*, Oxford University Press.
- Srinivasan, A. (2015). Are we lumious? *Philosophy and Phenomenological Research* .
- Stein, A. (2005). *Foundations of Evidence Law*. Oxford University Press.
- Thompson, J. (1986). Liability and individualized evidence. *Law and contemporary problems*.
- Wasserman, D. T. (1991). The morality of statistical proof and the risk of mistaken liability. *Cardozo Law Review*.
- Weinberg, J., Gonnerman, C., Buckner, C., & Alexander, J. (2010). Are philosophers expert intuiters? *Philosophical Psychology*.
- Wells, G. L. (1992). Naked statistical evidence of liability: is subjective probability enough? *Journal of personality and social psychology* .
- Williamson, T. (2000). *Knowledge and Its Limits*. Oxford: Oxford University Press.
- Yang, D. (2019). What's wrong with the modal conceptions of luck and risk. *Erkenntnis*.