

ARTICLE

Recent work on the proof paradox

Lewis D. Ross 

Department of Philosophy, Logic and Scientific Method, London School of Economics and Political Science, London, UK

Correspondence

Lewis D. Ross, Lakatos Building, Portugal Street, Holborn, London, WC2A 2HE, UK.
Email: l.ross2@lse.ac.uk

Abstract

Recent years have seen fresh impetus brought to debates about the proper role of statistical evidence in the law. Recent work largely centers on a set of puzzles known as the “proof paradox.” While these puzzles may initially seem academic, they have important ramifications for the law: raising key conceptual questions about legal proof, and practical questions about DNA evidence. This article introduces the proof paradox, why we should care about it, and new work attempting to resolve it.

1 | WHAT IS THE PROOF PARADOX AND WHY CARE?

A long-running debate concerns the proper role of statistical evidence in the law. This debate has important conceptual and practical ramifications for the administration of justice. Recent years have seen fresh impetus brought to this debate, with focus largely centered on a set of puzzles known as the “proof paradox.” This article introduces the proof paradox, why we should care about it, and new work attempting to resolve it.

There seems to be a difference between evidence presented in a statistical form (e.g., “her work patterns suggest an 80% chance that she was in the building”) versus non-statistical evidence (e.g., eye-witness testimony; CCTV recordings; confessions).¹ The proof paradox begins from the thought that deciding a legal case on the basis of *statistical evidence alone* can seem problematic. To clarify what exactly is meant by deciding a case on the basis of statistics alone, I need to say something about the requirements of legal proof.

For a legal case to succeed, various discrete claims must be established—each on the applicable standard of proof (i.e., rules that determine how strong a body of evidence must be before a court should treat evidence as establishing a claim). For example, in the criminal law, it is necessary to establish “beyond a reasonable doubt” both that the suspect performed the criminal *actus reus* (roughly, proscribed action) and possesses the requisite *mens rea* (roughly, blameworthy mental state like intention or recklessness). In the civil law of torts it is necessary to establish, on the “balance of probabilities,” that the pursuing party has been *harmed*, that the defending party *caused* the harm, was *negligent*, and had a relevant *duty of care* to the pursuing party. A case is settled on the basis of purely statistical evidence when, with respect to one of these essential claims, the only supporting evidence is statistical.

Consider now the following examples. The first is a civil case (in which only statistics support causal responsibility for harm) and the second is a criminal case (in which only statistics suggest that a proscribed action has been carried out):

BLUE BUS: A bus causes injury to a pedestrian, but it is not known which company the bus belongs to. On the route where the accident occurred, the *Blue Bus Company* runs 75% of the buses. There is no further information available to settle which company the bus belongs to. [Adapted from Tribe, 1971]

RIOT: An electronics store is struck by looters during a riot. On the day the riot occurs, 100 televisions are taken from the store: the transaction record indicates that only one was purchased legitimately. No receipt was issued. Joel is stopped by the police while carrying a television. Joel concedes he has one of the 100 televisions taken from the store—99 of which were stolen—but maintains his innocence. [Adapted from Smith, 2020]

According to a common view, despite the weighty statistical evidence, it would be inappropriate to either find the *Blue Bus Company* civilly liable or Joel criminally guilty in these cases.² Indeed, it has been suggested that rejecting bare statistics is legally routine, although this claim is not entirely uncontroversial.³ This reluctance to rely on statistics is a striking fact in need of explanation.

A chief reason that such cases have been thought “paradoxical” is that they seem in tension with the standards of proof used in the law. Recall the civil “balance of probabilities” standard (sometimes called “preponderance of the evidence”). In **BLUE BUS** it seems indisputable that the evidence against the *Blue Bus Company* makes their responsibility for harm more probable than not. Recall the criminal “beyond a reasonable doubt” standard. In **RIOT**, one might think that it is reasonable to doubt Joel’s guilt. However a 1% chance of error is extremely low—and we routinely base criminal convictions on fallible evidence that, on reflection, we would not always estimate as 99% reliable. And, even if you do not find 99% likelihood convincing beyond a reasonable doubt, it is easy to construct cases involving even more impressive statistics with an even lower chance of error. It is difficult to square reluctance to ever base legal decisions on bare statistics with the standards of proof found in the law. This raises questions about the very nature of legal proof.

Notwithstanding this important conceptual issue, a further reaction might be that the proof paradox lacks practical importance. This would be a mistake. Firstly, there are close real-life analogues to the **BLUE BUS** case.⁴ And secondly, the rise of DNA profiling compels a solution to puzzles surrounding statistical evidence.

It may surprise some readers to learn that, when evaluating DNA evidence, the court has *only a statistical estimate* presented by a forensic scientist to rely on.⁵ This estimate concerns how likely it is that an incriminating sample belongs to the suspect. The reason DNA gives rise only to a statistical estimate rather than an outright assertion of a match is that DNA profiles are fallible in the following sense: there is always a tiny chance that the observed similarity is merely fortuitous, that is, that there has been a *random match* between the incriminating sample and the DNA of somebody entirely unrelated. DNA evidence can lead to both *false negatives* and *false positives*. The relevance to the proof paradox is that (“cold-hit”) DNA evidence can sometimes be the *only evidence* linking the accused to the crime in question rather than vindicating pre-existing suspicion, as is brought out by the following example:

DNA: Someone is sexually assaulted in a secluded park. They cannot provide an account of the attacker’s appearance. DNA evidence from the crime matches that of someone on file for some unrelated reason. A forensic scientist estimates the chance of the DNA not belonging to that person to be 1 in ten million. [Adapted from Ross 2020]

There is not legal consensus on how to treat such cases. Working out the correct response is an important practical task facing contemporary legal systems that will affect the lives of those involved in criminal trials.

In practical terms, we can react in three ways to the proof paradox:

- basing a legal decision on bare statistics is *always* acceptable;
- basing a legal decision on bare statistics is *never* acceptable;

- basing a legal decision on bare statistics is only *sometimes* acceptable.

It is implausible to suppose that to rely on bare statistics is *always* acceptable. This is because there are often specific types of statistical evidence that would raise moral concerns if relied upon. Demographic statistics, for example, about race or sex, provide one example. Relying on such statistics might constitute an oppressive pattern against certain groups (Mogensen, 2019) or expose them to an unfairly heightened risk of mistaken conviction (Di Bello & O'Neil, 2020). Rather, the interesting debate concerns whether relying on bare statistics is *never* acceptable or is *sometimes* acceptable.

The former position, which to date has been more popular, could be vindicated by identifying some generic problem afflicting all cases relying on bare statistics. Such an approach would further suggest a unified explanation for our aversion to statistical evidence in both criminal and civil law. It is an open question whether this is a desideratum for responding to the proof paradox. Notwithstanding the popularity of such views, some have explicitly doubted whether a universal prohibition of bare statistics is desirable. For example, Di Bello (2019a) offers a decision-theoretic argument to the effect that certain cases (e.g., those involving 1/100 chance of error) but not others (e.g., those involving DNA evidence) will violate the aim of maximising utility while satisfying various fairness constraints; and Ross (2019) asks whether bare statistics could be more defensible in certain contexts (e.g., prosecuting sexual crimes, which suffer from perennially low conviction rates) than others. The acceptability of bare statistics in some cases but not others is another issue on which there is not currently consensus.

2 | THE PROOF PARADOX AND LEGAL PROBABILISM

A natural elucidation of the standards of proof is that they require establishing a claim to a given probability. For example, it is compelling to suppose that the “balance of probabilities” standard requires establishing a claim to a $>.5 p$. (The criminal standard would require a higher threshold: e.g., something in the region of $>.9 p$). Recent work calls this idea “legal probabilism.”⁶ The proof paradox is in tension with legal probabilism. One response is to use this tension to motivate a wide-ranging reconceptualisation of legal proof that vindicates aversion to bare statistics.

Cheng (2012) argues against the simple probabilistic account of legal proof. Rather, he argues, legal proof should be considered to be fundamentally *comparative*. Under this approach, the relevant comparison is whether the pursuer's story is *more likely* than the defender's story (and vice versa)—not a measurement of absolute probabilities.^{7,8} If we adopt this approach, the question to ask becomes: does the evidence observed makes either the pursuer or defender's narrative more likely? Framed this way, according to Cheng, bare statistics tend to be much less convincing. Take a concrete case: in **BLUE BUS**, the statistical evidence is equally compatible with a narrative on which the bus which caused harm was a Blue Bus or one on which it was some other type of bus. Hence, on Cheng's view, the statistical evidence does not suffice to make a substantial difference to the comparative likelihood of either story.

Other approaches invoke qualitative rather than quantitative revisions to legal proof. Gardiner suggests understanding legal standards of proof in terms of the ruling out of *relevant alternatives*, that is, “those error possibilities that cannot be properly ignored” (Gardiner, 2019:294). On her view, the different standards of proof correspond to more or less demanding criteria for deciding when an error possibility is relevant and must be excluded by the evidence. This conception of proof has interesting ramifications for the use of bare statistics, as there a strong sense in which no matter how statistically probable liability/criminal guilt may be in the **BLUE BUS** and **RIOT** cases, there will remain competing exculpatory narratives that cannot properly be ignored and which are not ruled out by the statistical evidence.

While some suggest that considerations relating to the proof paradox alone are enough to undermine legal probabilism,⁹ reconceptualising legal proof has ramifications beyond the use of statistical evidence. Moreover, it is possible to exclude bare statistics without a fundamental reconceptualization of legal proof. For example, one might argue that statistical evidence should not be admissible in court.¹⁰ Another possibility is to introduce a rule which

deems bare statistics “legally insufficient” evidence.¹¹ On each of these proposals, the rejection of bare statistical evidence would be carried out *independently and prior* to weighing it against a given standard of proof: so, we can reject bare statistics while preserving legal probabilism. As a result, whether or not to accept non-probabilistic interpretations of legal proof depends, in part, on debates further afield in legal theory (e.g., see Hedden & Colyvan, 2019 for a summary of such issues, and a defence of legal probabilism).

I now turn to approaches that focus specifically on bare statistics.

3 | EPISTEMOLOGISTS VERSUS STATISTICAL EVIDENCE

Epistemology studies the normative categories—for example, justification, knowledge and rationality—used to evaluate beliefs. It is no surprise, therefore, that epistemologists have had much to say about basing *beliefs* on statistical evidence. Many theorists, building on this work, attempt to pinpoint epistemic deficiencies in the courtroom use of bare statistics.¹² A general question for these approaches concerns the relevance of differences between belief-formation and legal verdicts. For example, Ross (2020) points out that individuals facing statistical evidence can *form and act on the basis of* hedged probabilistic beliefs (e.g., *S* is probably guilty), while courts can issue and act only upon binary verdicts (i.e., treat *S* as guilty without qualification or absolve *S* entirely of guilt). Putting this disanalogy to one side, I now survey a variety of approaches drawing on epistemological theory to address the proof paradox.¹³

3.1 | Sensitivity and safety

Many epistemological concepts invoke modal ideas—considerations to do with possibility and what might have happened. One well-known modal condition (e.g., defended as a condition for a belief to count as knowledge) appeals to a notion called *sensitivity*.¹⁴

The idea of sensitivity concerns tracking the truth. Beliefs can be true without being *sensitive* to the truth. For example: if I believe it is raining outside based on drawing tarot-cards, then my belief would not be sensitive to the truth even if true. Roughly, a belief which is sensitive to the truth is one we would not have held had that belief been false. It is easy to see how, in broad strokes, this idea is supposed to apply to statistical evidence.¹⁵ The thought is that verdicts supported only by statistics are not sensitive. Regardless of whether the Blue Bus caused the accident or not, the evidence against the *Blue Bus Company* would remain unchanged. We might suppose, regarding sensitivity, bare statistics stand in contrast with other types of evidence like eye-witness testimony.¹⁶

Safety conditions are a popular successor to sensitivity views in epistemology.¹⁷ In contrast to sensitivity views, which focus on what would have happened had the judgement in question been false, safety views instead focus on how easily one could have been wrong.¹⁸ A chief motivation for safety views is to preclude certain types of *luck* in epistemology—on an orthodox view luckily believing the truth blocks knowledge, even if the evidence seems otherwise convincing (e.g., you cannot *know* the time by looking at a stopped clock precisely when it shows the correct time). Epistemologists often precisify safety views by appealing to the idea that we can rank counterfactual possibilities (or “worlds”) with respect to how “near” they are to the actual world. Safe beliefs are those that are true in the nearby worlds. In this vein it has been suggested that bare statistics make it *too easy* for a verdict based on such evidence to be mistaken: bare statistics, so the thought goes, do not provide safe verdicts because error is always too near a possibility.¹⁹

An instructive observation about both views, one that probes our desiderata for attractive legal principles, is that they are extremely strict about when it is acceptable for a court to rely on evidence. Consider: it is plausible to suppose that a court can issue a false verdict (e.g., find someone guilty when they are in fact innocent) but still have made the right judgement relative to the available evidence. For instance, if many compelling eyewitnesses all convincingly testify that *x*, and there is no countervailing evidence against *x*, then arguably a court is entitled—and

indeed should—return a judgement that x . This is so even if x is false; if, improbably, each eyewitness had been mistaken, hallucinating, or dishonest.

Sensitivity views preclude acceptable false verdicts, because a legal verdict is never sensitive when false.²⁰ Recall: sensitive verdicts are ones that would not have been endorsed had they been false. But in cases involving acceptable false verdicts, the verdict *is* false, and *has* been endorsed.²¹ So, sensitivity views seemingly do not allow for acceptable false verdicts. On its face, it may seem like safety views can allow for acceptable false verdicts. Perhaps a body of evidence E can support a given proposition p in such a way that p 's falsity is not an easy possibility, even though p is in fact false? This is not straightforward to vindicate on an orthodox conception of safety. Safety views typically *never* allow that a false belief can be safe (because it is usual to stipulate that what is *actually* the case is always a near possibility, at least as near as any counterfactual possibility). So, if court verdicts are like beliefs in this respect, then false verdicts can never be safe. More work needs to be done both to see if this conclusion is inevitable and whether allowing for acceptable false verdicts is indeed a desiderata for any solution to the proof paradox.

3.2 | Normalcy

Although it is impossible to demand that courts aim for certainty, the proof paradox raises doubts about whether probability alone suffices. In a body of work, Martin Smith has developed a concept of justification that he claims “lies between probability and certainty” and uses it to diagnose the proof paradox.²²

Smith's account is centered on a notion called “normic support.” Normic support can be informally introduced with the idea of something “calling for explanation.” According to this theory, you are normically justified in holding some belief p only if the falsity of p , given your evidence, would call for some special explanation. For example: while it might be very *unlikely* that I have won the lottery based on probability alone, winning would not be *abnormal* in the sense of requiring special explanation—sometimes, your number just gets drawn. Hence, my belief that I have *lost* the lottery would lack normic support; its falsity would not call for special explanation. On the other hand, were I to look at my ticket and see that my numbers do not match the winning numbers, then my belief that I have lost *would* be normically supported. Some special explanation would be needed to explain why my belief supported by sensory evidence was mistaken. Smith observes that something similar seems to apply in many cases involving bare statistical evidence. Take the **BLUE BUS** case. While it might be somewhat unlikely for the bus that caused the accident to belong to another company, it would not call for any special explanation. Thus, on Smith's account, the mere fact that 75% of the buses are owned by one company would not normically support the proposition that a Blue Bus caused the accident. This suggests a general solution to the proof paradox, namely that we should only find a party liable or guilty on the basis of evidence that provides normic support to that verdict.²³

Smith's view has it that DNA evidence by itself also fails to provide normic support (it is well-recognised that sometimes random matches occur—it theoretically requires no special explanation). Di Bello (2019b) has worried that this approach will overgeneralise and exclude relying on other types of evidence (e.g., fingerprints) that are equally susceptible to well-theorised types of error. The worry is that Smith's account is overly strict.²⁴

3.3 | Knowledge

A burgeoning area of epistemological research is working out what *normative role* different epistemic states play. Various philosophers think that knowledge provides the “norm” for many important practices, for example, believing something, asserting something and treating something as a reason for acting.²⁵ According to these “knowledge-norms,” there is a sense in which one should not do these things unless one has knowledge. Take assertion. The idea is that you should not assert a proposition (e.g., “it's raining outside”) unless you know it.

Knowledge-norms have been used to explain the apparent impropriety of relying on bare statistics, because it is a fairly widely held assumption that bare statistics often fail to yield knowledge (e.g., in **RIOT** we do not *know* Joel stole the TV, even though it is very likely).²⁶ Blome-Tillmann (2017a) develops an argument against purely statistical evidence from the assumption that knowledge provides the norm for *action*; that is, one should not act as if p (e.g., act as if someone is guilty by punishing them) unless one knows that p .²⁷ However, he supposes that is too demanding to require that courts *must* have knowledge before issuing a verdict. Attempting to remedy this worry, Blome-Tillmann suggests that courts should only issue verdicts on the basis of evidence that *raises the probability* that they have knowledge.

One might have some doubts about Blome-Tillmann's view about probability here. For, even if mere statistics might never be sufficient by themselves to yield knowledge, we may think that statistical evidence can still make it more likely that a proposition is known, by providing some justification. This worry may not be fatal: a knowledge-centric approach may nonetheless support a legal rule on which we should not license acting on evidence which by its nature is unable to undergird knowledge of a party's liability or guilt.

3.4 | Belief & Credence

Another important area of research in epistemology concerns the relationship between *belief* and *credence*. Belief is a familiar concept, but we can also adopt credences—which are something like degrees of confidence. These degrees of confidence are compatible with the absence of a “full belief” because we can judge some proposition p “.6 likely” or “reasonably probable” despite not believing p outright.

The relationship between belief and credence is a much-vexed epistemological question. One relevant idea is that belief and credence play different roles. Buchak (2014) and Littlejohn (2018) suggest that only full beliefs, and not credences, can license blame: i.e., we should blame (and feel resentful) only if we *believe* someone did wrong, but not if we merely have some degree of confidence. From this view, it is possible to conclude that courts should only blame and punish those accused of crimes on the basis of evidence that is enough to support a full belief (and not merely a credence) that they are guilty. Moss (2018: Chapter 10) worries that these views struggle to generalise to the civil domain. While it may seem reasonable to suppose that criminal convictions require something like a full belief in guilt, this is harder to square with the civil balance of probabilities standard.

Moss's own solution appeals to a form of knowledge, but of a specifically probabilistic kind: she pioneers the idea that not only can we know outright propositions (i.e., S knows that p) but can also possess *probabilistic knowledge* (e.g., S knows that *probably* p , or S knows that it is .6 likely that p). As this gloss suggests, probabilistic knowledge can come in different strengths. This allows Moss to say that the civil standard requires *knowing* that it is $>.5$ likely that the defendant is liable, while leaving open exactly what strength of probabilistic knowledge is required to satisfy the criminal standard. Moss's probabilistic knowledge functions like “regular” outright knowledge insofar as possessing it requires ruling out certain possibilities of error.²⁸ This enables her to argue that we do not possess probabilistic knowledge in many proof-paradoxical cases.

4 | MORALITY, JUSTICE AND THE AIMS OF THE LAW

The previous section outlined attempts to identify an epistemic deficiency in statistical evidence. However, this raises important questions about the aims of the law. Why should the justice system care about anything other than accuracy? As Enoch et al. (2012) put it, imagine a legal system that complies with some epistemic constraint (e.g., dealing only in safe verdicts) and compare it with another system that is more accurate without these constraints. Why not prefer the more accurate system? Why countenance *accuracy sacrifices* in order to eschew statistics?

A variety of considerations concerning morality, justice and institutional rationality may help us answer this question, for instance by showing that these considerations align with the epistemic standards identified earlier.²⁹ Morally inflected solutions to the proof paradox are not entirely new. For example, several decades ago Wasserman worried that relying on bare statistical evidence “ignores the defendant's capacity to diverge from his associates or from his past, thereby demeaning his individuality and autonomy” (Wasserman 1992:942–3).³⁰ This approach has been further developed—and further challenges acknowledged—by Pundik (2008). However, until very recently such approaches have been overshadowed by discussions focusing on epistemology.

Enoch et al. answer the accuracy sacrifice challenge by appealing to *incentives*.³¹ Part of the role of a legal system is to incentivise and disincentivise different types of behaviour. Consider **RIOT**-type cases in which someone is wondering whether to pay for a television, or cases in which a bus company with dominant market share is wondering whether to improve safety standards. If these parties know that there will be incriminating statistical evidence against them *either way*, then the law will not provide the right incentives: why improve safety standards if you will be on the hook in Blue Bus-type cases regardless? Enoch et al. suggest that such considerations generate an argument for requiring evidence to be sensitive. While they make an important point about incentives, this line of thought will not generate a universal prohibition on bare statistics. For instance, in cold-hit DNA cases, individuals are not faced with the prospect of knowingly being “on the hook” regardless of how they act. A more specific legal concern is that defending parties can bring exculpatory evidence: in the **RIOT** case, a plausible alternative incentive would be to acquire proof of purchase; in **BLUE BUS** another incentive would be to better track one's fleet of buses (e.g., using GPS).

Some proof-paradoxical cases are structured so that relying on statistical evidence would *guarantee* erroneously holding someone responsible. For instance, suppose in the **RIOT** case we had captured 100 people carrying TVs, with only 1 TV purchased legitimately. A clear worry is, if we suppose that incriminating statistics suffice to hold a generic individual responsible, then it should license holding *every* TV-carrying individual responsible. But this would lead necessarily to erroneously convicting someone. Similar cases can appear in civil contexts.³² Some have used this thought to object to bare statistics. For instance, Nunn (2015), a legal scholar, argues that relying on such evidence violates a defending party's right to *due process*. This approach extends the legal prohibition on presenting multiple *incompatible* incriminating theories, where one theory by necessity must be false (Nunn, 2015:1418–1,421 for discussion).³³ Something to note about such approaches is that they will not lead to a general prohibition on relying on bare statistics. Necessarily blaming an innocent is not an essential property of proof-paradoxical cases, as illustrated by the **BLUE BUS** and cold-hit DNA cases.

These arguments use institutional and moral considerations to argue against the use of bare statistics. However, it is crucial to recognise that there are similar arguments which militate in the opposite direction, sometimes *favouring* the use of bare statistics.³⁴

This is best appreciated with examples. As we have discussed, tort cases require proving a causal link between the defender's activity and the harm—and, generally speaking, statistics alone have been thought not to establish such a link. However, there are certain harms concerning which it is particularly difficult to establish any causal link. Some legal systems have taken the view that we might want to relax the requirements of proof, relying instead on bare statistics, to advance the pursuit of justice when we have suboptimal evidence available. A first example, from UK law, concerns asbestos exposure. Asbestos exposure is linked to serious diseases, but establishing a causal link between any particular period of exposure and contraction of illness is difficult—victims are often exposed to asbestos over many years, by different employers, and in different contexts. As a result, UK courts have taken the view that statistical evidence about the risk of contracting an asbestos-related disease from a period of exposure to asbestos can suffice to hold an employer responsible for such a disease—even in lieu of evidence showing any causal connection.³⁵ A second example is the US doctrine of *market-share liability*. There are cases where consumers suffer harm after using a product—in the flagship case, a drug that turned out to be carcinogenic³⁶—but cannot identify which of a number of competing manufacturers produced the harm-causing product. Some courts, recognising the fact that refusing to rely on such statistics would doom to failure any case against these manufacturers, take the view that it is more just to apportion liability according to market-share in such scenarios. The reader will notice that there is a close structural resemblance between this case and the famous **BLUE BUS** scenario.

5 | CONCLUSION

Recent work on the proof paradox is an excellent example of what philosophy can contribute to legal theory, and why technical work in epistemology is relevant for questions of public importance. However, thinking about the role of the courts also forces us to examine the limitations of using theories designed to deal with individuals to make prescriptions about public institutions, prompts a discussion of what the aims of institutions like courts are in the first place, and of how best we should balance them when they conflict. As the use of new technology becomes an ever more prominent source of evidence, we will continue to face difficult questions about the proper role of statistical evidence. These questions will be best answered by using the expertise of experts from a variety of disciplines, including legal theorists, epistemologists, and moral philosophers.

FURTHER READING

Gardiner (2018) provides a helpful overview of different proof-paradoxical cases and a critical survey of several influential approaches.

Roth (2010) provides a survey of legal approaches and relevant case-law, offering a corrective to those who claim that statistical evidence is clearly in bad legal standing.

Pardo (2019) provides an alternative overview, focusing helpfully on how the proof paradoxes are related to other philosophical paradoxes.

Redmayne (2008) provides a helpful longer introduction, discussing some older views and drawing on relevant case-law.

ORCID

Lewis D. Ross  <https://orcid.org/0000-0001-9683-9924>

ENDNOTES

¹Of course, there may be a sense in which *any* evidence can be evaluated probabilistically. For example eye-witness evidence seems paradigmatically non-statistical, but we could statistically evaluate the chance of a given eye-witness being mistaken by comparing a reference-class of testifiers. The literature lacks a clear statement of the essential nature of statistical evidence, beyond appealing to way that evidence is presented in court. Future work would do well to address this issue.

²The reluctance to ascribe liability on the basis of statistics is known as the *Wells Effect* after an empirical study of mock-jurors (Wells, 1992). This is a piece with a broader phenomenon, often illustrated with *lotteries*, whereby subjects are reluctant to treat statistical evidence on par with other types of evidence (see Ebert, Smith, & Durbach, 2018).

³For example, Blome-Tillmann (2017) calls rejecting statistics the “conservative” view. For dissent see Roth (2010), Ross (2020), or Krauss (2019).

⁴See *Smith v Rapid Transit Inc.*, 317 Mass. 469, 470, 58 N.E. 2d.

⁵Here's a quote of the type of expert testimony that constitutes incriminating evidence in DNA cases.

Expert Witness: “We give an estimation of the chance of somebody randomly unrelated in the population, having the same profile, given that [the suspect has] the profile, is one in a thousand million.” [*Wilson v DPP* 2017 IESC 54 at 5.18.]

⁶The term appears at least as far back as Haack (2014).

⁷Formally, Cheng's increased likelihood requirement applied to the civil “balance of probabilities” standard can be stated as: find for the pursuer iff $\Pr(E/H^1)/\Pr(E/H^2) > 1$ where E is the evidence adduced, H^1 is the pursuer's story, and H^2 is the defender's story.

⁸There are similarities with Sullivan (2019) who proposes a likelihood-based theory of *legal fact-finding*, contrasting probabilistic (primarily Bayesian) and belief-based views of when evidence supports a given fact. Sullivan explicitly discusses bare statistics at (2019:45–53).

⁹For example, Smith (2020) is sympathetic to this view.

¹⁰For example, see comments in Blome-Tillmann (2017).

¹¹Evidence can be admissible without being sufficient; when evidence is insufficient, the case is thrown out before being judged against the standard of proof. See Ross (2020) for discussion.

¹²Many views are indebted to Thomson's (1986) discussion of the absence of the appropriate *causal relationship* between statistical evidence and verdict. For critical discussion see, for example, Gardiner (2018).

¹³Here, we focus on applying theories from epistemology to law. Backes (2019) discusses the other direction: using legal cases as evidence for views in epistemology.

¹⁴Originators of this view include Dretske (1970); Nozick (1981). Becker and Black (2012) for a recent collection.

¹⁵Enoch, Spectre, and Fisher (2012) influentially defend a version of this view; but, as it sometimes overlooked, their full view does not only appeal to sensitivity but further looks at how different approaches to evidence law incentivise and discourage certain types of conduct.

¹⁶See Blome-Tillmann (2015), Gardiner (2018), Smith (2018), and Pardo (2018) for criticisms of sensitivity views.

¹⁷Influential defenders of safety include Sosa (1999), Williamson (2000), Pritchard (2009).

¹⁸The difference between the two approaches is significant: it delivers conflicting judgements on the epistemology of necessary truths, and on whether knowledge is closed under known entailment.

¹⁹See Pritchard (2015, 2018) and Pardo (2018) for defences of safety views applied to the law, and critical comparison with sensitivity views.

²⁰Sensitivity views also seem to face a similar issue about ruling out *unwittingly* insensitive evidence (e.g., imagine a compelling eye-witness of the Blue Bus incident who was, unbeknownst to all parties, colour-blind). However, see Enoch and Spectre (2019) for a response, and defence of sensitivity views over safety views.

²¹One strategy for defenders of sensitivity here might appeal to generalisations about different types of evidence—for example, claiming that we should eschew statistics because doing so makes sensitive judgements generally more likely.

²²For the most comprehensive statement see his (2016) book. See Smith (2018) for discussion of the proof paradox.

²³Smith's view seems to allow for acceptable false verdicts. See Backes (2019: fn 19) for further discussion.

²⁴See Blome-Tillmann (2017b) for further critical discussion of Smith's view.

²⁵Williamson (2000) is the classic reference for many of these views, but his discussion is difficult. See Benton (<https://www.iep.utm.edu/kn-norms/#H4>) for an accessible overview.

²⁶See Papineau (2020) for scepticism about the preoccupation with knowledge in epistemology and with respect to statistics in the law.

²⁷Another suggestion is that court judgements are a species of *assertion* (e.g., see Thomson, 1986; Bolinger, 2018). If correct, this is another way to vindicate using epistemic norms to critique bare statistics.

²⁸This is a similarity with Gardiner's view of proof, introduced earlier.

²⁹See Redmayne (2008: IV) for critical discussion of older views appealing to moral considerations.

³⁰See Pundik (2008) for valuable discussion of this view, refining it and outlining some challenges.

³¹See Enoch and Fisher (2015) for a legally-oriented development of this idea.

³²A famous example, due to Cohen (1977), is the **GATECRASHER** case: suppose the majority but not all attendees at a rodeo hopped the fence and the organiser wants to recoup expenses. It is probable that any given individual is a gatecrasher. Yet, if we applied such reasoning to every member of the audience, we would be sure to hold liable an honest ticket-buyer.

³³Both Littlejohn (2018) and Smith (2020) discuss similar considerations, drawing out their connection to issues in epistemology concerning what it is rational to believe.

³⁴For a full defence of this claim, see Ross (ms).

³⁵See *Fairchild v Glenhaven Funeral Services Ltd* [2002] UKHL 22; *Barker v Corus UK Ltd* [2006] UKHL 20; *Sienkiewicz v Greif* [2011] UKSC 10 for notable developments of this doctrine.

³⁶The classic case remains *Sindell v. Abbott Laboratories*, 26 Cal. 3d 588.

REFERENCES

- Backes, M. (2019). Epistemology and the law: Why there is no epistemic mileage in legal cases. *Philosophical Studies*. <https://doi.org/10.1007/s11098-019-01337-6>
- Becker, K., & Black, T. (Eds.). (2012). *The sensitivity principle in epistemology*. Cambridge, UK: Cambridge University Press.
- Blome-Tillmann, M. (2015). Sensitivity, causality, and statistical evidence in courts of law. *Thought*, 4(2), 102–112.

- Blome-Tillmann, Michael (2017a). 'More Likely Than Not'–Knowledge First and the Role of Statistical Evidence in Courts of Law.
- Blome-Tillmann, M. (2017b). Statistical evidence, normalcy, and the gatecrasher paradox. *Mind*. In A. Carter, E. Gordon, & B. Jarvis (Eds.), *Knowledge first - approaches in epistemology and mind* (pp. 278–292). Oxford, England: Oxford University Press.
- Bolinger, R. J. (2018). The rational impermissibility of accepting racial generalizations. *Synthese*, 1–17. <https://doi.org/10.1007/s11229-018-1809-5>
- Buchak, L. (2014). Belief, credence, and norms. *Philosophical Studies*, 169(2), 1–27.
- Cheng, E. K. (2012). Reconceptualizing the burden of proof. *Yale LJ*, 122, 1254.
- Cohen, J. L. (1977). *The probable and the provable*. Oxford: Oxford University Press.
- Di Bello, M. (2019a). Trial by statistics: Is a high probability of guilt enough to convict? *Mind*, 128, 1045–1084.
- Di Bello, M. (2019b). Proof paradoxes and Normic support: Socializing or relativizing? *Mind*. <https://doi.org/10.1093/mind/fzz021>
- Di Bello, M., & O'Neil, C. (2020). Profile evidence, fairness, and the risks of mistaken convictions. *Ethics*, 130(2), 147–178.
- Dretske, F. (1970). Epistemic Operators. *Journal of Philosophy*, 67(24), 1007–1023.
- Ebert, P. A., Smith, M., & Durbach, I. (2018). Lottery judgments: A philosophical and experimental study. *Philosophical Psychology*, 31(1), 110–138.
- Enoch, D., Spectre, L., & Fisher, T. (2015). Sense and 'Sensitivity': Epistemic and Instrumental Approaches to Statistical Evidence. *Stanford Law Review*, 67:557. Sensitivity, Safety, and the Law: Reply to Pardo.
- Enoch, D., Spectre, L., & Fisher, T. (2012). Statistical evidence, sensitivity, and the legal value of knowledge. *Philosophy and Public Affairs*, 40(3), 197–224.
- Enoch, D., Spectre, L., & Fisher, T. (2019). *Legal Theory*, 25, 178–199.
- Gardiner, G. (2018). Legal burdens of proof and statistical evidence. In D. Coady & J. Chase (Eds.), *The Routledge handbook of applied epistemology*. London, England: Routledge.
- Gardiner, G. (2019). The reasonable and the relevant: Legal standards of proof. *Philosophy & Public Affairs*, 47, 288–318.
- Haack, S. (2014). *Evidence matters: Science, proof, and truth in the law*. New York: Cambridge University Press.
- Hedden, B., & Colyvan, M. (2019). Legal Probabilism: A qualified Defence. *Journal of Political Philosophy*, 27, 448–468.
- Krauss, S. (2019). Against the alleged insufficiency of statistical evidence. *Florida State University Law Review*, 47, 2020.
- Littlejohn, C. (2018). Truth, knowledge, and the standard of proof in criminal law. *Synthese*, 1–34. <https://doi.org/10.1007/s11229-017-1608-4>
- Mogensen, A. (2019). Racial profiling and cumulative injustice. *Philosophy and Phenomenological Research*, 98(2), 452–477.
- Moss, S. (2018). *Probabilistic knowledge*. Oxford: Oxford University Press.
- Nozick, R. (1981). *Philosophical explanations*. Cambridge, MA: Harvard University Press.
- Nunn, G. A. (2015). The incompatibility of due process and naked statistical evidence. *Vand. L. Rev.*, 68, 1407.
- Papineau, D. (2020). The disvalue of knowledge. *Synthese*, 1–22. <https://doi.org/10.1007/s11229-019-02405-4>
- Pardo, M. S. (2018). Safety vs. sensitivity: Possible worlds and the law of evidence. *Legal Theory*, 24(1), 50–75.
- Pardo, M. S. (2019). The paradoxes of legal proof: A critical guide. *Boston University Law Review*, Vol. 99, 2019; U of Alabama Legal Studies Research Paper No. 3293023.
- Pritchard, D. (2009). Safety-based epistemology: Whither now? *Journal of Philosophical Research*, 34, 33–45.
- Pritchard, D. (2015). Risk. *Metaphilosophy*, 46(3), 436–461.
- Pritchard, D. (2018). Legal risk, legal evidence and the arithmetic of criminal justice. *Jurisprudence*, 46(3), 436–461.
- Pundik, A. (2008). Statistical evidence and individual litigants: A reconsideration of Wasserman's argument from autonomy. *The International Journal of Evidence & Proof*, 12(4), 303–324.
- Redmayne, M. (2008). Exploring the proof paradoxes. *Legal Theory*, 14(4), 281–309.
- Ross, L. D. (ms). Justice and Evidential Gaps. See lewisdylanross.com/papers for full draft.
- Ross, L. D. (2020). Rehabilitating statistical evidence. *Philos Phenomenol Res*. <https://doi.org/10.1111/phpr.12622>
- Roth, A. (2010). Safety in numbers - deciding when DNA alone is enough to convict. *NYU Law Review*, 85, 1130.
- Smith, M. (2016). *Between probability and certainty: What justifies belief*. England: Oxford University Press.
- Smith, M. (2018). When does evidence suffice for conviction? *Mind*, 127, 1193–1218. <https://doi.org/10.1093/mind/fzx026>
- Smith, M. (2020). Against Legal Probabilism. In J. Robson & Z. Hoskins (Eds.), *Truth and trial*. Abingdon, UK: Routledge.
- Sosa, E. (1999). How must knowledge be modally related to what is known? *Philosophical Topics*, 26(1&2), 373–384.
- Sullivan, S. P. (2019). A likelihood story: The theory of legal fact-finding. *University of Colorado Law Review*, 90, 1.
- Thomson, J. J. (1986). Liability and individualized evidence. *Law and Contemporary Problems*, 49(3), 199–219.
- Tribe, L. (1971). Trial by mathematics: Precision and ritual in the legal process. *Harvard Law Review April*, 84(6), 1329–1393.
- Wasserman, D. T. (1992). The morality of statistical proof and the risk of mistaken Liability'. *Cardozo Law Review*. 13, 935.

Wells, G. L. (1992). Naked statistical evidence of liability: Is subjective probability enough? *Journal of Personality and Social Psychology*, 62, 739–752.

Williamson, T. (2000). *Knowledge and its limits*. Oxford, England: Oxford University Press.

AUTHOR BIOGRAPHY

Lewis Ross is Fellow of Philosophy and Public Policy at the London School of Economics. His research interests lie at the intersection between political philosophy, epistemology and legal philosophy.

How to cite this article: Ross LD. Recent work on the proof paradox. *Philosophy Compass*. 2020;15:e12667.

<https://doi.org/10.1111/phc3.12667>