



Cambridge  
Elements

Philosophy of Law

# The Philosophy of Legal Proof

Lewis Ross

# Contents

Introduction	1
1 Standards of Proof	1
2 Proof: Fixed or Flexible?	18
3 Should Proof Be Binary?	27
4 Legal Probabilism and Anti-Probabilism	36
5 Who Should Decide?	52
References	67

~~One might have a system where each of these receives different treatment: in terms of the duration of formal punishment, availability of appeal, possibility of parole, or severity of the punishment. Indeed, we noted earlier that some think that the death penalty should only be available when the crime is proven beyond any doubt. Implementing such a process across the entire criminal justice system would obviously require a radical rethink of criminal justice. I don't defend or advocate for this system, but I leave it as an exercise for the reader to consider what (if anything) is wrong with it. This will clarify your thinking about fundamental questions concerning the purpose of criminal justice and punishment.~~

## 4 Legal Probabilism and Anti-Probabilism

Proof is fundamentally about strength of evidence. Evidence makes accusations and claims more or less *likely*, can *explain* why something happened, can lead us to *believe* something or reject it, can render certain doubts *reasonable* or *unreasonable*. We have used this terminology more or less interchangeably so far. Now, I want to ask whether we can be more precise in understanding the relationship between evidence and proof.

### 4.1 Probabilism and Anti-Probabilism

Consider the following idea:

**Legal probabilism:** Legal proof is justifying the *probability* of guilt/liability above some threshold.

Probabilism, as the name suggests, views the standards of proof in terms of probabilities.<sup>59</sup> Probabilities are quantitative measurements of how likely something is, on a scale that ranges from 0–1. Something that has a probability of 1 is certain, 0 is certain not to happen, while something that has a probability of 0.5 is just as likely as not to happen. (If you prefer, you can convert these into percentages: e.g. 0.5 = 50 per cent.)

Probabilism seems to offer a pretty compelling diagnosis of the civil standard of proof. What is it to prove something on the ‘balance of probabilities’? Well, according to probabilism, it is to show that it is >0.5 likely to be true. Probabilism might seem less obvious when applied to the criminal standard of proof, that is, the BRD test. We might think that deciding whether a doubt is ‘reasonable’ is not just a matter of estimating the likelihood of error, but something more qualitative in nature. But if what makes a doubt reasonable is

<sup>59</sup> See Urbaniak and Di Bello 2021 for general introduction. Hedden and Colyvan 2019 summarise and respond to objections to Probabilism.

not about how likely the accusation is to be wrong, then what else could it be about? The difficulty of answering this question can be used to support probabilism.

We need to distinguish between debates about (i) how the current system works and (ii) how legal systems should be designed. One question is – can we interpret ‘beyond reasonable doubt’ in a probabilistic way? This question is separable from the larger question – should we implement standards of proof defined explicitly in terms of probabilities? Defining criminal proof in terms of a precise probability (e.g. 0.9 likelihood of guilt) is one way to escape the notorious vagueness that many see in the current BRD standard. Some think the fact that BRD is so hard to define is problematic – after all, shouldn’t people know exactly what standard they will be judged against?<sup>60</sup> However, as discussed earlier, the vagueness of the criminal standard can be defended because it empowers the judge or jury to apply standards that are most appropriate to the case at hand.

What I do in this section is to try and make progress in understanding legal proof by articulating and explaining why some hold the opposing view to probabilism:

**Legal anti-probabilism:** Something cannot be legally proven just by showing that it exceeds some probability threshold.

Anti-probabilism is a negative claim, denying the truth of probabilism. It has been endorsed by a wide range of philosophers working on legal philosophy. On anti-probabilist views, one cannot prove something *simply* by showing that it is very likely. There are many different types of anti-probabilist view which have lots of detail and nuance. However, before looking at these, we should say something about why anyone might be sceptical about probabilism in the first place.

First, it might seem arbitrary to say that someone deserves punishment if their guilt is proven to a 0.95 likelihood but not a 0.94 likelihood. Can this be used as an argument against probabilism?<sup>61</sup> I don’t think so. We could use the same argument for a 0.94/0.93 probability and so on right down to 0. But this surely wouldn’t be right! There are sharp cut-offs under almost any way we think about proof. Even if we think only in qualitative terms, the difference between a reasonable and an unreasonable doubt will sometimes be determined by small differences in the evidence or psychology of the fact-finder. The probabilistic approach simply makes sharp cut-offs explicit. We also use sharp cut-offs in other

---

<sup>60</sup> For example, see [Laudan 2006](#).

<sup>61</sup> You might reflect on [Section 3](#) on intermediate verdicts and develop this objection.

areas of life (e.g. deciding who gets medical treatment) so it isn't obvious that the existence of sharp boundaries is objectionable when difficult decisions must be made.

Another worry could be that just focusing on probabilities omits other important notions concerning how evidence can support a conclusion. For example, we might be interested in the coherence of evidence or how well corroborated evidence is. Perhaps these notions are hard to capture in purely probabilistic terms? Evaluating this abstract challenge would require spending a long time discussing different ways of interpreting the idea of probability to see if our preferred notion could capture these other evaluative ideas.<sup>62</sup>

Since we are limited for space, I will take a different approach and focus instead on what has been the biggest challenge to probabilism in recent years – a purported type of counterexample. The challenge comes from the ‘proof paradox’.

## 4.2 The Proof Paradox

The proof paradox is generated by cases where a conclusion has strong probabilistic support yet many resist judging the conclusion proven. The cases used to generate the proof paradox involve *statistical evidence* supporting guilt or liability. The idea will be best appreciated by considering influential examples from the literature.<sup>63</sup>

### 4.2.1 Civil Law

**Gatecrasher.** The organisers of the local rodeo are suing John for gatecrashing their event. Their evidence is as follows: John attended the Sunday afternoon event– he was seen and photographed there. No tickets were issued, so John cannot be expected to prove he bought a ticket with a ticket stub. However, while 1,000 people were counted in the seats, only 300 paid for admission.<sup>64</sup>

**Blue Bus.** A bus negligently 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 per cent of the buses. There is no further information available to settle which company the bus belongs to.<sup>65</sup>

<sup>62</sup> See section 2 of Hedden and Colyvan 2019 for a flavour of how this debate might go.

<sup>63</sup> For a different summary, see Redmayne 2008 or Ross 2020.

<sup>64</sup> Adapted from Blome-Tillmann 2020, 565; original case due to Cohen 1977.

<sup>65</sup> Adapted from Tribe 1971, 1340–1.



### 4.2.2 Criminal Law

**Prisoners.** One hundred prisoners are exercising in the prison yard. Ninety-nine of them suddenly join in a planned attack on a prison guard; the hundredth prisoner plays no part. There is no evidence available to show who joined in and who did not.<sup>66</sup>

**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.<sup>67</sup>

### 4.2.3 Discussion

While these examples are philosophical inventions there are real legal cases that resemble these stylised scenarios.<sup>68</sup> And, as I later show, there are other cases with huge public significance that are structurally similar.

Many feel uncomfortable about relying on purely statistical evidence to convict someone of a crime or hold them liable for a civil wrong. However, this discomfort is not easy to reconcile with the standards of proof. The civil law provides the clearest illustration. The civil standard of proof is ‘the balance of probabilities’. It seems, by stipulation, more likely than not that the Blue Bus caused the accident or that a generic rodeo attendee is more likely a gatecrasher than not. The criminal cases have a less paradoxical flavour – one might think it reasonable to have doubts in both the Riot and Prisoners cases. But why? After all, the probability of guilt is 99 per cent. Surely *other* criminal cases are settled on weaker evidence, evidence that would not support such a high confidence in guilt. For example, there is well-documented evidence on the limitations and unreliability of eyewitnesses in stressful situations.<sup>69</sup> Many eyewitness accounts might not be judged 99 per cent reliable. Yet, absent exculpatory evidence, the evidence of a direct eyewitness is sometimes regarded as sufficient for conviction. So, it remains puzzling why strong statistical evidence cannot perform the same role.<sup>70</sup>

We need to distinguish two questions: the psychological question of how people tend to react to statistical evidence, and the normative question of how the legal system *should* react. There is, I think, a clear psychological difference in how people react to statistical versus more direct types of evidence. Indeed,

<sup>66</sup> Adapted from Redmayne 2008, 282–3. <sup>67</sup> Adapted from Smith 2020, 93.

<sup>68</sup> *Smith v. Rapid Transit Inc.* [317 Mass. 469, 58 N.e.2d 754], for instance, resembles the Blue Bus case.

<sup>69</sup> For example, see Loftus 1996.

<sup>70</sup> The focus is on whether it is acceptable to rely on statistical evidence alone to convict or hold someone liable. The debate is not primarily about whether statistical evidence should be admissible, nor whether statistical evidence can be exculpatory.

discomfort about relying on purely statistical evidence is known as the ‘Wells effect’, after the psychologist – Gary Wells – who published a paper describing it.<sup>71</sup> The question is whether this psychological reaction is an irrational anti-statistical bias or whether it is justified.

Many have tried to vindicate the anti-statistical intuitions generated by the proof paradox. If this can be done, we have a counterexample to probabilism. These discussions have a long history, the proof paradox having been debated intermittently by legal scholars since the mid twentieth century, before being rediscovered more recently by philosophers.<sup>72</sup> I want to start by discussing recent work primarily carried out by philosophers working in epistemology.

### 4.3 Epistemic Responses to the Proof Paradox

Philosophers have noted that the proof paradox has important similarities with famous epistemological puzzles, most notably puzzles about lotteries.

Some propositions about lotteries are stupendously likely to be true. Consider the proposition ‘any given ticket in a ten-million ticket lottery is a losing ticket’. Despite being overwhelmingly likely to be true, many philosophers think that such propositions, based on probabilities alone, are different from other propositions we regularly rely upon.<sup>73</sup> It’s been popular to suppose, for instance, that we don’t *know* that we have lost the lottery just by reflecting on how unlikely winning is. This is puzzling, because there are *many* things we take ourselves to know even though we presumably have more than a one-in-ten-million chance of being wrong. For example, you might know you will attend a meeting later, even though occasionally meetings get cancelled unexpectedly – and surely more frequently than one-in- ten-million meetings! If we want to avoid conceding that the scope of our knowledge is much more limited than usually supposed, there must be some difference between the probabilistic evidence we have about the lottery and evidence for regular things that we do know.

Something to note about ‘lottery beliefs’ is that, even though they are very likely to be correct, there is another sense in which they are not secure. They are insecure because the evidence supporting these beliefs is completely compatible with your belief about impending lottery loss being false. Suppose you *did* have the winning ticket. Nothing would be different from your perspective. All the same probabilistic evidence against your victory would remain. Perhaps *non*-statistical evidence provides a tighter or more direct connection with the

<sup>71</sup> See Wells 1992.

<sup>72</sup> Classic earlier work on the proof paradox includes Cohen 1977 and Tribe 1971.

<sup>73</sup> See Ebert, Smith, and Durbach 2018 for an empirical study of lottery propositions and philosophical references.

truth? This is an imprecise, inchoate thought. An influential way of responding to the proof paradox tries to develop it.

The general strategy has been to borrow ideas from epistemology about the rationality of belief. The assumption behind this strategy draws on something discussed in [Section 1](#) – that legal verdicts and individual beliefs should be judged against similar standards. Of course, we have already noted that acquittals can be legitimate even if you don't fully believe in the innocence of the accused. Rather, the idea is that *guilty verdicts* should be based on evidence that would make believing in guilt rational. We can put the idea like this:

**Legal doxasticism.**<sup>74</sup> Guilty verdicts must only be issued when it is rational, given the evidence, to believe that the accused is guilty.

But what are the requirements for a belief to be rational? I will briefly introduce four theories that flesh out this idea.

*Sensitivity.* A belief is sensitive when it has the following property: it is a belief you would not have had if the belief were false. Some methods tend to produce sensitive beliefs (e.g. using your eyes in good conditions, because the fact you believe an object is there depends on the object being there) while other methods do not produce sensitive beliefs (e.g. predicting the future using tarot cards, because the future isn't determined by the cards you draw). Sensitivity is thus a way to capture the intuitive idea of 'tracking the truth'. Beliefs formed on the basis of statistics alone are not sensitive. Take the lottery case again – if you trust the statistics, you would have formed the same (false) belief about losing even in the event that you had a winning ticket. Some argue that we should only convict people when our belief in their guilt is sensitive – we acquire these beliefs by using methods that track the truth.<sup>75</sup>

*Safety.* Safety is a property of belief, often informally presented as a belief that couldn't easily have been false. This might sound uninformative, but the idea is usually developed by appealing to the idea that we can rank possibilities (or 'worlds') with respect to how close or far they are from the actual world. In this way, you can compare counterfactual situations against each other in terms of how close a possibility they were – for example, a world where humans are 200 feet tall is further away than a world where it rained yesterday morning. Safe beliefs are beliefs that are true in all nearby worlds. Some claim that beliefs formed on the basis of mere statistics are often unsafe because statistical evidence doesn't show that an incompatible conclusion was counterfactually

<sup>74</sup> 'Doxastic' means pertaining to belief.

<sup>75</sup> Sensitivity in epistemology is famously defended by Robert Nozick as a condition on knowledge. See [Melchior 2019](#), chapter 2 for an introduction. It is defended in legal philosophy by (among others) [Enoch, Spectre, and Fisher 2012](#) and [Enoch and Spectre 2019](#).



distant. For example, you might say that winning the lottery is always a close possibility even though statistically unlikely; all that has to happen for you to win is for some balls to drop into a hopper differently. Some argue that we should only convict when our belief that someone is guilty is safe – when the evidence shows that the accused being innocent is only a distant possibility.<sup>76</sup>

*Knowledge.* Knowledge is widely considered to have deep cognitive and social importance. Although (almost) everyone agrees that knowing something entails that it is true, how to further analyse the nature of knowledge is a topic of perennial debate. (Some of the properties discussed above are candidate requirements for a belief to count as knowledge.) It's usually taken for granted that relying on certain types of statistical evidence – as in the lottery – doesn't provide you with knowledge. One idea, popular in some quarters, is that knowledge provides the 'norm' for various practices: forming beliefs, making assertions, relying on something when deciding how to act. For example, the 'knowledge norm of assertion' dictates that you shouldn't go around telling people something will happen unless you know it will happen. Some have used this type of reasoning about the centrality of knowledge to various practices to argue that knowledge is the 'norm' for criminal conviction – you shouldn't pronounce someone guilty unless you know they are guilty (which you don't when relying only on statistics).<sup>77</sup>

Each of these views faces a common problem. They appear too demanding as conditions for a positive legal verdict. Suppose a court convicts someone after diligently seeking out and weighing the evidence, relying on the latest theories, consulting the most relevant witnesses, after an impeccable police investigation, and with the defence lawyer putting up a robust fight. Alas, it turns out the witnesses got it wrong (suppose they all made an honest mistake) and the person was innocent. Still, courts can only judge on the basis of the evidence they have. There will be cases where it is rational for a court to return a verdict that is (unbeknownst to everyone) false. If all the evidence points towards guilt, then it is rational for a jury to convict. If false verdicts can sometimes be rational, given extremely misleading evidence, none of the aforementioned properties can account for this. You can't know the verdict is false (you can't know anything false). False verdicts can also never be sensitive – if the person is innocent and your belief was sensitive then you wouldn't believe they are guilty. Likewise, safety views also seem too strict. This is because philosophers usually stipulate that the actual

<sup>76</sup> Safety in epistemology is famously defended by [Sosa 1999](#), [Williamson 2000](#), and [Pritchard 2009](#). It is defended in legal philosophy by (among others) [Pardo 2018](#) and [Pritchard 2022](#).

<sup>77</sup> Knowledge conditions in legal philosophy are defended by (among others) [Blome-Tillmann 2017](#), [Littlejohn 2020](#), and [Moss 2022](#).

world – that is, what actually happens – is *at least as near a possibility* as any counterfactual possibility. So, false beliefs can never be safe.

Of course, defenders of these views have various responses to this criticism. For example, one might focus on the *tendency of different methods* to produce beliefs that are safe, sensitive, or known – rather than focusing on whether the individual verdicts have these properties. This would of course require spelling out how good the tendency needs to be to pass muster, as well as an account of how to individuate methods.<sup>78</sup> However, we are limited for space so rather than get into these technicalities I want to forge on to another proposed epistemic condition on conviction.<sup>79</sup>

*Normalcy.* Normalcy is a type of non-probabilistic justification.<sup>80</sup> Informally, normalcy can be introduced with the idea of something ‘calling for special explanation’. Some propositions can be highly improbable while being, in the explanatory sense, entirely normal. Winning the lottery is an example; we don’t require a special explanation when someone wins the lottery, regardless of how improbable it may be. However, when an eyewitness says ‘I watched her steal the money!’, we would typically demand a special explanation upon learning that the eyewitness was mistaken and the accused actually innocent. We would demand a special explanation for why the accused was innocent, given that the eyewitness claimed to see them commit the crime. The normic view deals with proof-paradox cases by pointing out that it would *not* take special explanation for belief based merely on statistical evidence to be false. For example, in the Blue Bus case, it would not take special explanation for it to have been a Red Bus causing the accident – after all, the Red Bus Company runs 25 per cent of the buses on the route.

An advantage of ‘normic’ theories is that a belief can be justified in the normic sense despite being false. For example, if all the eyewitness evidence seemed to point towards guilt, it *would* require special explanation if it turned out that the accused was innocent. Hence, the false belief in guilt would have normic justification. The normic view seems to escape the ‘demandingness’ worry afflicting other epistemic diagnoses of the proof paradox.

But the normic view has its own problems.<sup>81</sup> One is about the idea of something requiring special explanation. You might worry that this idea is relative to: (i) how well informed we are about the fallibility of the evidence, and (ii) the way in which the evidence is presented. For example, as our

<sup>78</sup> This has proven a notoriously difficult project in epistemology! See [Conee and Feldman 1998](#) for a discussion of how hard it is to individuate methods for forming beliefs.

<sup>79</sup> There are also other epistemic conditions that have been defended as conditions on legal proof. For example, see [Gardiner 2019](#) and [Lackey 2021](#).

<sup>80</sup> Normalcy in defended in epistemology by [Smith 2016](#) and in legal epistemology by [Smith 2018](#).

<sup>81</sup> For further discussion of objections to normic views, see [Di Bello 2020](#).

psychological knowledge grows, the less surprising it becomes for an eyewitness to get things wrong. An expert in psychology might not demand a special explanation for the fallibility of an eyewitness, since they know eyewitnesses get things wrong all the time, especially in stressful situations. Indeed, the expert might know eyewitnesses simply get things wrong  $x$  per cent of the time. One day, we might even have eyewitness accounts presented in court in the form of a probabilistic estimate as to how likely they are to be accurate. (Indeed, a deep philosophical question looms – is there any underlying truth about whether evidence is *really* probabilistic or non-probabilistic, or does it rather depend just on the way that the evidence is presented?) What is normically supported might shrink in the face of our advancing knowledge. Yet, it strains our credulity to suppose that we shouldn't convict people if there is both statistical and eyewitness evidence against them – even if neither is normically supported in the eyes of someone who perfectly understands the fallibility of such evidence.

A more general objection to all epistemic accounts is that intuitions about evidence that lacks the aforementioned properties seem less secure when we *combine* different types of probabilistic evidence. Consider the following variation on the Blue Bus case.

**Blue Bus variation:** 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 Company runs 75 per cent of the buses and the Red Company 25 per cent of the buses. Fresh tyre marks are found at the scene of the accident that an investigator's uncontested report states were caused by the offending vehicle. All parties agree these could only be made by a certain brand of bus tyre. A recent insurance application form shows that 90 per cent of Blue Company buses have the implicated brand of bus tyre, while only 5 per cent of Red Company buses do. Moreover, police find a bus hubcap on the road immediately after the crash. Only 2 per cent of Red Company buses were recorded as having the implicated brand of hubcap, while 96 per cent of Blue Company buses have it.<sup>82</sup>

At the end of the day, the totality of the evidence against the Blue Bus Company remains statistical. Yet, it seems less compelling to suppose that we should not hold them responsible.

#### 4.4 DNA Evidence

I want to discuss a final interesting test case for probabilism – convictions based on 'cold-hit' DNA evidence.<sup>83</sup>

<sup>82</sup> Taken from Ross 2021b.

<sup>83</sup> DNA evidence is further discussed in Roth 2010 and Ross 2021a.

Everyone knows that forensic evidence can play an important role in criminal investigations. For instance, DNA evidence is often regarded as a gold standard in linking somebody to an unsolved murder or sexual assault. However, people are generally less familiar with the nature of such evidence and how it is presented in court.

Perhaps surprisingly, DNA evidence is presented in court just as a probabilistic estimate. A forensic expert provides the evidence. But they do not categorically say ‘this DNA belongs to that person’. Rather, they make claims like ‘the probability of the DNA *not* belonging to that person is one in ten million’. Why do they hedge their bets? Because, for any apparent DNA match, there is the tiny possibility that it is a *random match*. In other words, it is possible for an incriminating sample to *seem to match your DNA* even though you had nothing to do with the crime. Often, this tiny risk is washed out because DNA evidence is usually combined with other evidence linking the accused to the crime. But, in so-called cold-hit cases, such as where new techniques allow recovery of DNA evidence about historic crimes, we might not have any evidence apart from the matching DNA sample.

What should we do if all we have is the DNA match? Courts have struggled to decide whether it is acceptable to convict someone if the only evidence is a matching DNA profile. On the one hand, DNA evidence gives rise to probabilities *much* greater than those found in regular proof-paradox cases, as strong (or stronger than) a one-in-ten-million chance of error. Rejecting cold-hit DNA evidence would make for a less accurate criminal justice system. On the other hand, there is something ‘lottery like’ in DNA profiles, given the possibility of a random match. This means we have an interesting way to test intuitions about the proof paradox and probabilism. The heart of cold-hit DNA evidence is just an incriminating statistic giving rise to a probability estimate of guilt.

Interestingly, this generalises to other types of forensic evidence. There have been attempts to estimate fingerprint matches in probabilistic terms, rather than relying on dubiously reliable qualitative assessments. There can also be improbable random matches between unrelated fingerprint samples, since there is always a tiny statistical chance that two people are born with practically indistinguishable fingerprints.

Pushing the argument further, we might then consider cases involving conjunctions of incriminating forensic evidence and probabilistic evidence.

**Prisoners and DNA.** One hundred prisoners are exercising in the prison yard. Extremely grainy CCTV footage shows that ninety-nine of them attack and kill the guard. The 100th prisoner played no role in the assault and could have done nothing to stop it. From the footage it is impossible to distinguish which prisoners were involved. The ninety-nine murderers escape in one

direction and, some time later, the 100th prisoner escapes in a different direction. One prisoner is recaptured. Upon testing, it is found that his DNA matches the most dominant DNA profile found on a discarded switchblade used in the murder. The forensic scientist estimates the chance of a random DNA match as one in ten million.<sup>84</sup>

Here, it seems rather far-fetched to claim that it would be inappropriate to convict. Yet, on one reading, the evidence remains purely statistical. So, perhaps the ‘proof-paradox’ argument against probabilism isn’t so decisive?

#### 4.5 Moral and Political Diagnoses

Some have worried that focusing only on epistemic ideas fails to see the woods for the trees. Enoch, Fisher, and Spectre argue that some legal philosophers are exhibiting a type of ‘epistemic fetish’ by focusing on epistemic properties rather than the moral–political values that, ultimately, the legal system is intended to promote.<sup>85</sup> If this is right, it would be better to think about these moral or political ends first, then see where the epistemology fits in.

Just as with epistemological diagnoses of the proof paradox, many moral and political approaches take an inchoate thought and try to develop it. The kernel at the heart of many accounts begins from the idea that justice requires treating people as *individuals* and that relying on statistical evidence somehow fails to do this. By assuming that someone is just the same as other members of a reference class (e.g. rodeo attendees, prisoners, rioters) a worry is that we somehow disrespect or degrade the individuality of the person being accused. This inchoate thought seems to chime with what is wrong with statistical evidence in other contexts. For example, another debate concerns demographic profiling – using statistical evidence about racial, ethnic, or gender groups to attribute properties to people.<sup>86</sup> In this debate, many think we somehow treat people badly by lumping them in with other members of their group.

Various accounts try to sharpen this idea. One is that basing legal verdicts on mere statistics *disrespects* the individual, while another is that it fails to show due regard for one’s *autonomy* to diverge from one’s peers.<sup>87</sup> Others have suggested that ignoring individuality shows that the state is insufficiently concerned with the individual’s *right to security* from being harmed by false convictions.<sup>88</sup> Such accounts can then be used to backwards engineer an account of why certain types of evidence – evidence that is somehow

<sup>84</sup> Taken from Ross 2021b. <sup>85</sup> Enoch, Fisher, and Spectre 2021

<sup>86</sup> For example, see Gendler 2011; Bolinger 2020; Ross 2022.

<sup>87</sup> See Levanon 2022 and Wasserman 1992, respectively. <sup>88</sup> See Adams 2023.

*individualised* – is morally important. Indeed, perhaps such ideas about the importance of individualised evidence could explain the moral importance of the various epistemic properties discussed earlier.<sup>89</sup>

One immediate issue with this individuality-based approach is that the defending party in the proof paradox need not be an individual at all.<sup>90</sup> As the Blue Bus case shows, the party inculpated by statistics can be a corporation. Indeed, the defending company can be a powerful multinational that has, in all likelihood, harmed a vulnerable individual. It's not obvious that we owe it to such corporations to respect their individuality in a way that rules out using statistical evidence against them. I return to this issue at length shortly.

Diagnoses that appeal to individuality and the importance of individualised evidence focus on what we owe to people, rather than the consequences of relying on statistical evidence. But some reject statistical evidence by reflecting on the effect that relying on such evidence might have.

One concern is that *if* we rely on statistical evidence, we might commit ourselves to a procedure that would guarantee eventually making a mistake.<sup>91</sup> For example, take the Gatecrasher case. If we decided the evidence was strong enough to sanction one person, then by parity of reasoning it would be strong enough to sanction everyone at the rodeo. But then we would be sure to sanction a large number of innocent people, which seems patently unjust. However, while an important point, this cannot be a *general* diagnosis of the proof paradox. In some cases – like the Blue Bus or DNA cases – we don't have any certainty that relying on statistics about a given incident would guarantee a mistake, nor are a large group of people put 'on the hook' by relying on statistics.

David Enoch, Levi Spectre, and Talia Fisher have argued that relying on bare statistics is no good given the *incentive structure* we want the law to create. Obviously, we want people to follow the law – and we want evidence law to give people an incentive to follow the law by making sure that, if they break the law, there will likely be admissible evidence that can be used against them. Perhaps relying on statistics can create perverse incentives? In some cases it might. For example, consider the decision-making of someone wondering whether to buy a ticket in the Gatecrasher case described earlier (recall: no record will be provided of their purchase). If they know that statistical evidence is enough to find them liable, then whether or not they purchase a ticket will make no difference to the inculpatory evidence against them. If we have a general principle that inculpatory evidence should be created by one's choices and

---

<sup>89</sup> For discussion, see [Mortini 2022](#). <sup>90</sup> This point is also made by [Pundik 2008](#).

<sup>91</sup> For example, see [Nunn 2015](#).



actions, then this is a mark against purely statistical evidence which remains inculpatory pretty much independently of what an individual chooses to do.<sup>92</sup>

A worry with this story about incentives, however, concerns whether it is supposed to be an empirical claim or just a theoretical one. Empirically, it isn't clear that proof-paradoxical cases are common enough to really have any substantial effect on the incentives created by the law. For example, do we really think that the mere possibility of Blue Bus-style cases influences a transportation company CEO? Perhaps the worry is more theoretical rather than practical. But, it isn't clear why this theoretical worry matters so much.

To explain why, I want to draw attention to an under-emphasised side of the debate: the moral problems with *refusing* to rely on statistical evidence. The existence of these moral issues not only casts doubt on the idea that relying on statistics is always wrong, but also calls into question whether it is sensible to look for a single general response that aims to capture every case of purely statistical evidence.

#### 4.6 The Cost of Denying Statistics

Classic 'proof-paradox' cases direct us to focus on the possibility of punishing the wrong person. However, this obscures the fact that refusing to rely on statistical evidence can be to the extreme detriment of the person *harmed*. Proof-paradox scenarios involve an 'epistemic gap' – a situation where we lack the knowledge needed to identify the party responsible for a harm we know has been wrongfully caused. Take the Blue Bus case. Something easy to overlook is that, if we *don't* rely on statistical evidence, we might leave a person who has been negligently squashed without compensation. These issues become sharper in cases where the probability of error is very small and where there is an economic imbalance between the harmed party and putative wrongdoer. For example, consider the following variant on the Blue Bus case.

**Monopoly Bus:** A bus negligently causes injury to a pedestrian, who is left with life-changing injuries and unable to work. But there was no eyewitness evidence linking the bus to a particular company. On the route where the accident occurred, only one outfit has a regular service: the Monopoly Bus Company. Uncontested statistics from analysing CCTV cameras in adjacent neighbourhoods show that only 1 in every 10,000 buses passing through that area is owned by a private individual. There is no further information.<sup>93</sup>

<sup>92</sup> Indeed, as Enoch et al. (2012) point out, this is a mark in favour of methods that produce *sensitive* verdicts – because sensitive methods are sensitive to individual choices and actions.

<sup>93</sup> Ross 2021c, 326.

The evidence remains statistical, but I feel no reluctance about relying on it. This shows that our judgements are sensitive to the relative position of the parties and various justice-related factors – not just to the epistemic properties of the evidence. Indeed, I think that this suggests that there is not necessarily anything paradoxical about (many) so-called proof-paradox cases. Sometimes we should rely on statistical evidence when doing so is to the benefit of the least well off.

One might complain that these cases are philosophical inventions or so rare as to be irrelevant to real legal practice. The complaint would be mistaken. Structurally similar issues have arisen in some of the most controversial legal cases, particularly in ‘tort’ law. Tort law is the branch of law that provides compensation for negligently caused harms, where being negligent does not reach the standard needed for something to count as a criminal matter. For example, if a shopkeeper’s sign is sloppily installed and hits you on the head, this may be a tort that you are entitled to be compensated for.

A normal requirement in tort law is that the person who is harmed shows that the other party *caused* the harm. For example, you might be asked to prove that the shopkeeper hung the sign, and their shoddy handiwork is the reason that it fell on your head. Causation can be easier or harder to prove depending on the case. But sometimes it is nigh impossible.

Litigation about asbestos is one important example.<sup>94</sup> As is well known, and known long before it was effectively regulated, asbestos causes serious and sometimes fatal damage to the human respiratory system. Industrial labourers – and their partners who cleaned their asbestos-caked clothes – got seriously ill and in many cases died premature deaths from asbestos exposure. This was typically because employers failed to take reasonable steps to protect workers from exposure, such as by providing adequate safety equipment. However, labourers often handled asbestos while working for many different employers during their lives. How could they prove that their illness was caused by any *particular* employer? Many of their employers were negligent (by failing to provide safety equipment) but none could be shown to be responsible for any particular illness. Just think about the impossibility of tracing the influence of some asbestos fibres on a disease and demonstrating that these fibres entered the body during one period of employment rather than another.

The way causation is usually understood in the law is through a ‘but for test’. In other words, you ask ‘but for  $x$ , would  $y$  have happened?’. If the answer is no, then  $x$  is said to have caused  $y$ . In asbestos cases, however, it was impossible to

---

<sup>94</sup> I discuss this at length in [Ross 2021c](#). Another important example is the doctrine of market-share liability. For additional discussion, see [Krauss 2020](#).

show that the disease wouldn't have happened but for the time the labourer spent working at any particular company. Yet, clearly all of the employers were at fault. Given that these labourers and their partners were sometimes suffering fatal illness due to employer negligence, it would clearly be a massive injustice if they were unable to gain compensation.

In such cases, relying on statistics is all the court can do. The only evidence the labourers could cite was statistical information about the estimated level of exposure each employer had been responsible for and epidemiological statistics about how likely a generic person was to contract a disease given a particular level of exposure. Framed this way, we see that the labourers were being forced to play a potentially fatal lottery by their employers. The courts took the view that such statistical evidence *could* suffice to assign liability to the employers for the diseases.<sup>95</sup> The general approach of holding the employers liable based on statistics seems right to me. In cases such as this, there is a strong case for thinking that any intuitive discomfort we have about relying on statistics is outweighed by imperatives of justice.

Beyond the fact that relying on statistics could remedy an injustice, are there deeper principles that justify using them in such cases? I think so. In the asbestos case, all the employers exhibited similar failings. Even if one employer was 'lucky' in that they didn't actually cause the disease, they were equally blameworthy as the similarly negligent employer that was 'unlucky' in causing the disease. One might prefer that liability is shared between parties in virtue of their failings, rather than distributed by the vagaries of chance. For example, consider the following principle:

**Shared standards:** Whenever a harm is negligently caused but falls into an epistemic gap, it is reasonable to apportion responsibility among potential harmers where they share similar standards when conducting the risky activity.<sup>96</sup>

This principle captures the idea that liability should follow negligent behaviour rather than the purely chance matter of whether the negligence actually causes harm.

Now, here is the even more controversial part. If we accept such a principle, then there may be situations where it is similarly justified to use statistics to attribute liability *even in the regular Blue Bus case*. After all, if the companies

<sup>95</sup> See *Fairchild v. Glenhaven Funeral Services* [2002] UKHL 22 16; *Barker v. Corus* [2006] 2 A. C. 572 17; *Sienkiewicz v. Greif* [2011] UKSC 10. There are different ways to apportion liability: one is to 'share' full liability, another is to make the parties each 'part' liable in proportion to the risk they have caused.

<sup>96</sup> Taken from *Ross 2021c*, 327.

each have similar standards, then they are each imposing risks on the public. And such corporations will generally be better able to bear the burden of financial loss as a cost of business compared to the individual who will go without compensation if denied the use of statistical evidence. The Blue Bus case is supposed to be one of the clearest examples of a case where we should not rely on statistics alone. But, once we take into account the different moral considerations, it is not altogether obvious that the traditional reaction to this case is correct.

#### 4.7 Statistics in Criminal Law Reconsidered

I have defended the use of statistics to fill epistemic gaps in the civil law, when using statistics averts serious injustice. This, to my mind, is evidence that we cannot reject probabilism across the board. But does similar reasoning apply to the criminal law?

Recall something we noted in passing when discussing DNA evidence. Rejecting statistical evidence can require that we accept accuracy sacrifices in the legal system. If we refuse to rely on cold-hit DNA evidence, we release people who are *overwhelmingly* likely to be guilty. Echoing the discussion of Larry Laudan in [Section 1](#), recall that if we fail to convict the guilty, this might mean that more crimes are committed as a result. Enoch, Fisher, and Spectre use this observation to pose the provocative question: ‘How many more people are you willing to have assaulted, or murdered, or raped under your designed system, just in order to secure [some epistemic status] for the findings of your criminal justice system?’<sup>97</sup>

For example, imposing a ‘sensitivity’ requirement on guilty verdicts would (arguably) mean we can’t convict some apparent sexual offenders on the basis of cold-hit DNA evidence, even though the likelihood of guilt is stupendously high. But why should a legal system do anything other than try to be as accurate as possible? What do we gain by caring about knowledge or other epistemic properties? The sharpest way to put the worry is: why should we willingly be *less* accurate, just in order to promote knowledge, sensitivity, safety, or normalcy?

This is a difficult question! My own answer would return to the idea outlined in [Section 1](#) about the centrality of belief to findings of criminal guilt.<sup>98</sup> Criminal law, as we have discussed, is distinctive compared to civil law; findings of criminal guilt lead to moral blame and the possibility of retributive punishment. It is important that guilty verdicts are based on the sort of evidence that can lead to members of the community believing the person is guilty.

---

<sup>97</sup> Enoch, Fisher, and Spectre 2021, 89–90.    <sup>98</sup> Also, see [Ross 2023a](#).

Aiming to convict only on the basis of evidence that makes it rational to believe something is the best way to make sure this happens. Statistical evidence often fails to generate a full belief; rather, it just elicits a probabilistic estimate. So, there is a natural argument for why courts might refuse to rely on mere statistics – mere statistics don't tend to support a full belief in guilt in the mind of the community. Still, this argument might not rule out statistics in every case. Perhaps DNA evidence, involving such tiny chances of error, does tend to elicit full belief in the guilt of the accused (compared to regular proof-paradox cases involving much shorter odds).

My general view is that discussions of the proof paradox should look at the details of the case at hand. There may be no single resolution to the question of whether we should rely on statistical evidence alone; rather, there will be some cases where it is acceptable and others where it is not. Whether we should be probabilists or anti-probabilists is a case-dependent matter and must be approached by looking at contextual considerations of justice and policy.

### ~~5 Who Should Decide?~~

~~We now turn to our final question: *who* should decide the outcome of a trial? What person or group should be trusted with deciding whether the standards of proof have been met, thus determining whether the accused is guilty or not?<sup>99</sup>~~

~~In some periods of history, communities seemed to leave the decision to God. 'Trial by combat' (letting the disputants fight it out) and 'trial by ordeal' (having the accused perform some risky or wounding task) were both seen as ways as testing the sincerity – the 'oath' – of those accused of wrongdoing. If their oath was good, according to the official story, God would intervene to ensure that they prevailed.<sup>100</sup>~~

~~But even in these times, communities were not content to entirely separate proof from the available evidence. Trial by ordeal, for example, was often ambiguous. One ordeal was to pluck a stone from a cauldron of boiling water. If the inevitable wound healed cleanly, it was a sign of innocence; if it festered, it was a sign of guilt. But determining whether a wound is on its way to healing cleanly is a matter of interpretation – one that must be made by humans, even those claiming to interpret on behalf of a supernatural entity. Given what local people knew about the evidence, this would influence their decision.~~

~~Legal systems today answer the 'who should decide' question in strikingly different ways. Some leave the decision entirely in the hands of a professional judge who makes judging their career. Others continue the now ancient practice~~

<sup>99</sup> Jurors are also occasionally used in civil trials, for example in assessing defamation cases.

<sup>100</sup> For example, see [Baker 2019](#).