Cambridge Elements

Philosophy of Law

The Philosophy of Legal Proof

Lewis Ross

https://doi.org/10.1017/9781009127745 Published online by Cambridge University Press

ISSN 2631-5815 (online)
ISSN 2631-5807 (print)

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

https://doi.org/10.1017/9781009127745 Published online by Cambridge University Press

~~decide particular trials, the standards of proof emerge. This 'bottom-up' way of thinking about proof entrusts a great deal of responsibility to those who decide cases, relying on their judgement about what the appropriate way to think about the case is. To return to the phrase with which we began this section, it would mean that each judge or jury has to make up its own mind about what is required for an accusation to be proved to a 'moral' certainty.~~

3 Should Proof Be Binary?

Legal proof is usually conceived as a binary. Consider the words of Lord Hoffmann, an influential English judge:

> 'If a legal rule requires a fact to be proved . . . a judge or jury must decide whether or not it happened. There is no room for a finding that it might have happened. The law operates a binary system in which the only values are 0 and 1. The fact either happened or it did not.'[44]

The first paragraphs of this Element were highly sympathetic to Hoffman's description. I suggested that courts cannot shrug their institutional shoulders and that 'maybe' was no use in legal adjudication. In criminal cases, which we continue to focus on, trials end with somebody being convicted or acquitted. Although we've discussed how criminal proof might be flexible by using stronger or weaker standards in different situations before convicting somebody, we haven't yet questioned the *binary nature* of legal decisions. However, there are deep questions about whether this binary system is the best way to approach legal proof.

3.1 Binary versus Non-Binary Systems

Binary legal systems predominate across the world. These are systems in which criminal trials end in either a single 'Guilty' verdict or a single verdict of acquittal (usually phrased as 'Not Guilty').

But not all jurisdictions use a binary system of proof. In Scotland, for example, a 'third verdict' called *not proven* has existed for several hundred years.[45] *Not proven* exists alongside *guilty* and *not guilty* as a separate option for the court. Not proven is the verdict returned when the judge or jury is not satisfied with returning a not guilty verdict, but also cannot find the accused guilty beyond a reasonable doubt. In this sense, not proven is what we can call an *intermediate verdict*. The not proven verdict officially leads to the same outcome as not guilty – the accused is released without punishment. However,

[44] *Re B* [2009] 1 AC, para. 2.

[45] There are plans to abolish this verdict, and at the time of writing it seems likely that it will be abolished.

https://doi.org/10.1017/9781009127745 Published online by Cambridge University Press

this does not mean that the *informal* consequences are the same – and, as we will explain, the verdict may not have the same social meaning. There is also an intermediate verdict in Israeli law, which has two types of acquittal; one is a 'regular' acquittal, while a second is an 'acquittal on the basis of doubt', where the doubt refers to doubt about the guilt of the accused. In both cases the accused is released, but doubtful acquittals can lead to diminished rights to compensation for false imprisonment, among other consequences.[46]

Legal history also describes non-binary systems of proof that have been abolished. Some of these are not as benign for the accused as the Israeli or Scottish systems. Federico Picinali describes systems with intermediate *punishments* to go along with intermediate verdicts, such as in the *ius commune* that predominated across Europe for centuries following the rediscovery of Roman law in the form of Justinian's Digest in the early 1000s.[47] This system permitted 'extraordinary punishments' that were less severe than the ordinary punishment for the crime. But they served up a stiff penalty for an intermediate confidence in guilt, as the penalties could include exile. Examples from the distant past aside, Italy retained a third verdict until near the end of the twentieth century, a verdict that allowed an official 'criminal record' to be kept regarding the acquitted person and could justify additional measures such as surveillance.

One could mistakenly suppose that intermediate verdicts are a mere curiosity. In fact, thinking about the strengths and weaknesses of different intermediate systems cuts to the heart of fundamental questions about the proper role of criminal justice. Despite these intermediate verdicts raising fascinating questions about legal philosophy, they are a surprisingly understudied phenomenon – both theoretically and empirically. As we will see, it is far from obvious that a binary system really is best.

3.2 Motivating Non-Binary Systems

To see the appeal of non-binary systems, we must think more carefully about what binary systems involve. There's a common misconception that courts either find people guilty or *innocent* of what they are accused of. But this is wrong. Technically, courts only find people *not guilty* when they decline to convict them. Now, this might sound exceptionally pedantic. But these ways of thinking about the court's role are crucially different.

Putting the point in abstract terms, saying that a standard for judging that x is not satisfied is very different from saying that the evidence supports not-x.

[46] For discussion, see Vaki and Rabin 2021.
[47] See Picinali 2022 for insightful discussion of all the systems mentioned in this paragraph.

https://doi.org/10.1017/9781009127745 Published online by Cambridge University Press

Consider an example. Suppose I plan to climb a mountain at the weekend. The mountain is dangerous, so I adopt the following standard: I will only climb if I think it is >90 per cent likely to stay dry. Now, if the weather forecast shows that this standard is not satisfied, this doesn't mean that it is 90 per cent likely to rain! Rather, it just means that it is at least 10.1 per cent likely to rain. The same is true about verdicts in criminal courts. Just because we don't find someone guilty after judging evidence against a very demanding standard of proof (e.g. 'convict only if you think their guilt has been proven beyond reasonable doubt'), this doesn't mean the evidence suggests that they are – or are likely to be – innocent. Finding an accused person not guilty is compatible with *a broad range of confidence in their guilt*; it only rules out the very highest levels of confidence in guilt. You can think that it is rather likely that someone is guilty while properly acquitting them.

Is it a problem that there is so much variation in how an audience can interpret a not guilty verdict? To answer this question, we might try to see things from the perspectives of two different parties affected by criminal justice.

> ***'The perspective of the accused'***. Suppose you are accused of a serious crime. You maintain your innocence. There is a trial – you are found not guilty. But being found not guilty is different from being found innocent, so you feel you have not managed to 'clear your name'. Employers and acquaintances view you with suspicion, thinking there can be no smoke without fire. All you can truthfully say is that the verdict demonstrates the court did not think your guilt had been proven beyond a reasonable doubt. Being found not guilty in a criminal court doesn't prevent people from treating you in ways that reflect their suspicion of you. For example, you might still lose your job in a school even if found not guilty of sexual misconduct, or you might not be hired for a banking job even if even if accused but not convicted of fraud.
>
> ***'The perspective of the community'***. Someone in our community has been accused of a serious crime and found not guilty. However, we don't know how to react to this finding. On the one hand, perhaps the person was simply the victim of mistaken identity and poses no risk to us. On the other hand, perhaps the person only narrowly escaped being convicted, because the evidence against them was strong but not quite strong enough to convince everyone on the jury. We are caught between the temptation to protect ourselves and a concern not to punish them for a crime they have not been convicted of. We waver between the two interpretations, neither fully reintegrating nor rejecting the accused.

Perhaps you care more about one of these perspectives. However, the two complaints are really the same – binary-verdict systems do not provide enough information. Trying to solve the problem of one perspective will also address the complaints of the other perspective.

https://doi.org/10.1017/9781009127745 Published online by Cambridge University Press

A non-binary system of proof would address these complaints by removing ambiguity in what an acquittal means. By giving the court additional options, we learn something extra about what the judge or jury made of the evidence. Consider the 'not proven' verdict again. It's a verdict that lies between guilty and not guilty. If the court returns a not proven verdict – interpreted as an intermediate between guilty and not guilty – we learn that they are *not* willing to say that the evidence strongly supports the guilt *or* the innocence of the accused.[48] The existence of additional verdicts also enriches the information provided by the 'regular' verdicts. For instance, if a court returns a 'not guilty' verdict in a system where they could instead have returned a not proven verdict, we also learn something; namely, the court has *declined* to use the not proven verdict and must be happy to say that the evidence supports a not guilty verdict. To summarise, systems with more than two verdicts have an *informational richness* lacked by binary systems. This means that someone found not guilty in a system with multiple verdicts might feel that they have better cleared their name – and the wider community can also feel more reassured about the weakness of the evidence supporting guilt.

Of course, once a legal system decides to abandon binary verdicts, this raises difficult questions about the rules that should be used for choosing a verdict. Having more than two verdicts means that it's possible that no specific verdict is a majority or even plurality winner among the jury, even if there is an odd number of jurors. Indeed, there can even be 'preference cycles' where no single verdict wins a Condorcet-style competition.[49] But these are not insuperable difficulties – the fact that there are existing and successful non-binary systems demonstrates that they do not create huge practical problems.

Intermediate verdicts are just one way to make criminal courts more expressive. There are other ways to speak to some of the same concerns, adding more information to trial decisions. Currently, criminal verdicts are not accompanied by any written justification. But throughout history, courts have not just provided verdicts of guilty and not guilty, but also given reasons or explanations for their decisions. This system of 'reasoned verdicts' gives the judge or jury the chance to express how confident they are in the innocence of the accused when they deliver an acquittal in the context of a regular binary-verdict system. This is an extremely interesting approach with its own particular strengths and

[48] An oddity of the Scottish system is that judges offer no guidance on when to return a not proven verdict. Therefore, my comments are a reconstruction of what the verdict should mean.

[49] Bray 2005 discusses preference cycles, alongside a helpful general discussion of the advantages of intermediate verdicts.

https://doi.org/10.1017/9781009127745 Published online by Cambridge University Press

weaknesses – I invite the reader to further consider these.[50] For now though, I continue to focus on non-binary systems of proof.

We have been considering the idea that the information-poor nature of binary systems leaves those acquitted open to social stigma and other adverse reactions. But perhaps we should see courts as having a more limited role, one that does not purport to care about social stigma? On a restricted conception of their role, courts *only* inquire into whether the evidence against the accused is strong enough to blame and punish them, taking no interest in social reactions to criminal proceedings beyond this. Indeed, one might have concerns about the idea that legitimising stigma – as might happen when an intermediate verdict is returned – is any business of the court. Court-sanctioned stigma, after all, is different from stigma that exists simply because private individuals put their own interpretation on a verdict. Perhaps by getting into the business of actively licensing social stigma, we invoke the extra worry that courts could actively harm the innocent in cases where an intermediate verdict licenses the stigma of someone who is truly innocent. Moreover, perhaps the state does not wish to facilitate employers or potential relationship partners from rejecting accused persons on the basis of accusations that haven't been proven on the high standard of criminal proof. (Although, when it comes to intimate relationships and susceptibility of partners to harm, one might well one might well think that excluding people who are quite likely to be guilty is perfectly reasonable.) These are valid concerns and much depends on how you view the proper role of the court. But we must bear in mind that harming the innocent via stigmatisation *already* occurs when we acquit an innocent person who is stigmatised due to the informational paucity of a binary system. This is something that most criminal justice systems enable, even though there are alternative systems available with multiple verdicts.

3.3 Worries about Non-Binary Systems

Let's consider whether we lose anything by trying to enrich the legal system with non-binary verdicts.

One important value we keep returning to is public confidence in criminal justice. We want citizens to trust courts. Might having more than two verdicts – going beyond the binary – make people trust the criminal justice system less? Perhaps having intermediate verdicts could be seen as publicly admitting that courts are not always fully confident in their decision to let someone go or to punish them, making transparent the fallibility of trials.

[50] For discussion and defence of reasoned verdicts, see Coen and Doak 2017.

https://doi.org/10.1017/9781009127745 Published online by Cambridge University Press

This is not, in my view, a decisive worry. Any sense of infallibility that courts enjoy is illusory – we should not cultivate the idea that the public resemble uneducated peasants trembling before the infallibility of a Leviathan. Public confidence in criminal justice should not be secured by *misunderstanding* the nature of the law and what it can reasonably achieve.

Another related reason to be suspicious of non-binary approaches is that they sound like a 'cop-out' – a failure to decide rather than decisively settling the case one way or the other. While this reaction is tempting, I am not convinced it is right. It may be easy to hear an intermediate level of confidence as indecision or an invitation to continue investigating the subject – that is, saying you haven't made your mind up *yet*. But this isn't a necessary part of what it means to adopt an intermediate level of confidence in something. Taking an intermediate or uncertain attitude about a question can be a fully rational and indeed definitive response to the available evidence. Suppose someone wants to make a bet over a coin flip (the coin is regular; a fair coin). Before flipping, I ask whether you believe you will win. The only rational answer is to say you are not sure – that your level of confidence is in the middle or that you are suspending judgement. Given the available evidence, and indeed *given all the evidence there could be* (unless you develop supernatural powers of foresight), this is the right thing to say. In fact, it is the only sensible thing to say. Moreover, having an intermediate level of confidence is not necessarily a prelude to further inquiry. If all the evidence has been gathered and it doesn't decisively point one way or the other, then remaining in an intermediate position forever is what you should do.[51]

Of course, when it comes to criminal trials the state *could* always gather more evidence: greater resources could be expended on policing; specialist detectives could be flown in; forensic data could be re-checked. But this is true of every case! A guilty or not guilty verdict in a binary system of proof is just a way of saying: 'here is the best judgement we can come up with, given the available evidence'. Court verdicts are always relative to the evidence that has been presented in court; they do not say anything about the completeness of the evidence, nor do they provide a cast-iron guarantee that we should not reopen the inquiry at a later time. Indeed, even if the evidence at one time strongly supports a given conclusion, this doesn't mean that new evidence cannot turn things on their head to support the opposite conclusion instead. Understood properly, there is not anything about an intermediate verdict that makes it 'less final' than classic binary verdicts.

However, there is a different way to push this idea of intermediate verdicts being a problematic 'compromise'. Some studies on the Scottish not proven

[51] Consider the religious agnostic!

https://doi.org/10.1017/9781009127745 Published online by Cambridge University Press

verdict have suggested that third verdicts serve to draw juries away from guilty verdicts.[52] You can imagine the process as one where one juror proposes to convict, a second juror raises some doubts, and all agree on a third juror's proposal to return a not proven verdict as a way to 'meet in the middle'. We can conjecture that the presence of an intermediate verdict reduces the number of convictions overall.

You can react to this 'compromise' phenomenon in different ways.[53] On the one hand, it could be seen as a good thing to prevent juries or judges from convicting when they are not entirely sure of guilt. Perhaps intermediate verdicts focus the mind by reminding the jury just how confident they should be before finding someone guilty. Indeed, the fact that jurors are split in their individual verdicts might be a good indication that they should collectively agree to 'split the difference' and endorse an intermediate verdict. It indicates that the evidence is not clear-cut in favour of guilt or innocence. On the other hand, it could be seen as a type of cowardice that allows the fact-finder to escape the unpleasant business of sticking their neck out and making a difficult decision. If intermediate verdicts short-circuit proper and full deliberation, then this would be a problem. An intermediate verdict agreed after careful deliberation is defensible; returning an intermediate verdict to escape deliberation is not.

Since empirical work suggests that including intermediate verdicts does reduce the number of convictions, it must also reduce the number of false convictions. But, some worry that this comes at the disproportionate cost of allowing too many guilty people to escape justice. Given the ongoing argument about low rates of conviction for sexual criminality, some have argued that mechanisms which reduce conviction rates are problematic. It *is* true that intermediate verdicts are more common in sexual offences than in other types of offence, something that campaign groups have been keen to highlight. For example, in their campaign against non-binary verdicts, Rape Crisis Scotland points to the fact that 'not proven' acquittals are more than twice as common in rape cases (44 per cent) compared to other offences (20 per cent), along with the fact that rape cases have the lowest conviction rate compared to any crime type.[54] This is a key reason why the Scottish Government seeks to abolish the not proven verdict and return to a binary system. (A similar trend is found in Israel, where 'acquittal on the basis of doubt' is more common in sexual offence cases than in other types of case.[55])

[52] See Chalmers, Leverick, and Munro 2022.

[53] Something worth considering is psychological research on 'extremeness aversion', where people tend to prefer – and be biased towards – options that lie between extremes. For a meta-analysis, see Neumann, Böckenholt, and Sinha 2016.

[54] See Rape Crisis Scotland n.d. [55] Rabin and Vaki 2023.

https://doi.org/10.1017/9781009127745 Published online by Cambridge University Press

An assumption at play in this argument is that the 'base rate' of guilty people is much higher than currently reflected in conviction rates for sexual offences – namely, that it is worth making the trade-off of a few more innocent people being convicted in order to convict many more guilty people. Even if this is right, there are further questions we might have about this argument. One, for instance, is whether it is acceptable to remove the safeguard that intermediate verdicts provide in *all* crimes just to make it easier to get convictions in sexual offence trials. Another is whether it is legitimate to make assumptions about what a 'better' conviction rate would be, without being seen to endorse the problematic idea that public authorities should have targets regarding the proportion of convictions.[56]

3.4 Radical Departures from the Binary

We've restricted ourselves to discussing systems where there are only three verdicts, with two being different types of acquittal. However, there are many ways in which legal systems could offer even more precise evaluations about the strength of evidence. In theory, we could have very information-rich systems, where courts return precise probability estimates about the guilt of the accused.[57] There is something dystopian-sounding about this, but it is hard to put a finger on what. We might doubt that humans are especially good at thinking with mathematical precision about probability of guilt. It has a ring of artifice to firmly distinguish evidence supporting a 60 per cent versus a 70 per cent chance of guilt. Perhaps more fundamentally, ternary systems (like 'not proven') are still rooted in a sort of everyday practice; there is nothing unusual about remaining neutral or having an intermediate confidence about something after considering all the evidence. But precise probability estimates depart considerably from the ordinary way in which we evaluate each other. Still, it is worth raising such systems as a possibility, even if only because asking why we don't do things this way sheds light on our philosophical commitments.

Even if systems assigning precise probabilities to guilt are unpalatable, there are various other approaches that attempt to incorporate the idea that we should have a wider menu of possible responses to criminal accusations. For example, take the ternary 'not proven' system. One could suppose that it is appropriate to bar someone with a not proven verdict for sexual assault from certain types of employment, while not barring the same person if they receive a 'full' acquittal. There are many variants here. Federico Picinali, for example, discusses what he

[56] See Thomas 2023 for important evidence that suggests that juries are not primarily responsible for the low conviction rate for sexual offences.

[57] For a helpful discussion, see Spottswood 2021.

https://doi.org/10.1017/9781009127745 Published online by Cambridge University Press

calls a 'conditional acquittal'. This would involve giving a *harsher* punishment to the person given the conditional acquittal if found guilty at a later date of another offence.

In this provocative vein, I want to close this section with an even more radical way of proportioning the reaction of the court to the evidence: the possibility of graded punishment for the guilty. In everyday life, sometimes the best thing is to proportion our practical reaction to the evidence. For example, sticking with mountain-related examples, think about how weather can affect holiday walking plans. If it's certain to rain, you might call off the trip (no point trudging in the wet and cold). If it's certain to be sunny, you might plan to walk on a particular day (e.g. Saturday – then you can do something different on Sunday, like read legal history). If the weather is less predictable, you might plan to stick around in the area for longer (say, Saturday *and* Sunday) to give you best chance to complete the hike. It's often sensible to change how we act in cases where the evidence isn't decisive, doing something different from what you would do if the evidence were decisive in either direction. So, why not do the same in criminal justice? Why not punish the guilty in proportion to how confident we are in their guilt?

Under the most comprehensive version of a graded punishment system, we would punish most harshly those whom we were 100 per cent confident are guilty, punish slightly less harshly those whom we were only 95 per cent confident were guilty, and so on, continuing to lower the level of punishment as our confidence in guilt decreased. Presumably, there would be some level – some 'blame line' – below which we wouldn't punish at all. (It would be perverse to punish someone we thought was only 1 per cent likely to be guilty!) While at first blush such a system may sound perverse, it is not entirely unmotivated. For example, with any conviction, there is the risk of making a mistake. If a mistake has been made, any retributive punishment is (arguably) immoral, any rehabilitative efforts are wasted, and the public expense of punishment has been for naught. Proportioning the level of punishment to the level of confidence in guilt is a way to hedge our bets.

Even putting aside systems that use very fine-grained confidence levels, there are more coarse-grained alternatives available. Talia Fisher has explained that one might have a system involving multiple guilty verdicts corresponding to different standards of proof.[58] To illustrate, consider:

- Guilt beyond any doubt.
- Guilt beyond a reasonable doubt.
- Guilty on the basis of clear and convincing evidence.

[58] Fisher 2021.

https://doi.org/10.1017/9781009127745 Published online by Cambridge University Press

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.~~[59] ~~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~~

[59] See Urbaniak and Di Bello 2021 for general introduction. Hedden and Colyvan 2019 summarise and respond to objections to Probabilism.

https://doi.org/10.1017/9781009127745 Published online by Cambridge University Press