Introduction

Criminal courts can remove the liberty and even the lives of those accused of wrongdoing. Civil courts can deport asylum seekers, render companies employing thousands of people bankrupt, and remove children from the care of their parents. Using the right standards when deciding legal cases is therefore of the utmost importance in making sure those affected by trials receive a fair deal.

While legal standards can at first sound like a rather dry or technical subject, the project of deciding on the right standards raises fascinating and deep philosophical questions. These questions cut to the heart of debates about ethics, politics, psychology, and epistemology. Questions about legal standards force us to examine where the state's duty to protect society conflicts with the interests of those accused of wrongdoing. Understanding these debates also reveals how citizens can limit or control legal institutions that wield considerable – and potentially oppressive – power over them. Thinking about legal standards of proof is both a topic of great theoretical interest and one that affects the lives of everyone in society. The standards we use ultimately determine when the state can take away our freedom, our children, our property, and even our lives.

This Element is an introduction to the philosophy of legal proof. It aims to be accessible to students of both law and philosophy, presupposing no technical background in either subject.

The Element is organised around five questions.

- Section 1 introduces the standards of proof and asks what justifies them.
- Section 2 asks whether we should use different standards in different cases.
- Section 3 discusses whether criminal trials should end in binary outcomes – guilty versus not guilty – or whether we should use more fine-grained verdicts.
- Section 4 asks whether proof is simply about showing that something is probable or likely, concentrating on the famous 'Proof Paradox'.
- Section 5 considers who should be trusted with deciding the outcome of trials, focusing on the debate surrounding the jury system.

1 Standards of Proof

Societies use trials to resolve disagreements. These disagreements can be between private individuals, between corporations, and with the state itself. The disagreements can concern any number of issues, from mundane questions about who started a drunken fist fight, to disputes about the arcana of shipping law, questions about the results of an election, or criminal responsibility for murder or rape. Some disagreements end with trivial resolutions like trimming

a garden hedge found to be encroaching a neighbour's property. Other disputes end with outcomes of great severity, including the bankruptcy of corporations that employ thousands of people, the overturning of elections, or even the imposition of the death penalty.

Courts must be decisive when resolving disagreements. They must form a view on what happened and deliver a judgement about what should happen next. A court cannot end its work with a 'maybe' and a judge cannot throw up their hands and say they cannot decide one way or the other. Those accused of crimes must be punished or released; election results must stand or be overturned; at-risk children must be removed from the family home or kept where they are. This burden of deciding is difficult because courts are usually confronted with ambiguity. Trial participants often fundamentally disagree – this is why there is a trial! – and point to seemingly contradictory evidence.

Since courts must decide one way or another, they need some method for dealing with evidence and opinions that point in different directions. Courts therefore rely on various rules concerning when something should be taken as 'proven'. While questions about legal proof may sound dry or technical, they are in fact of central importance in legal and political philosophy. Think about it this way: the rules we choose to govern trials are really rules about when the state should use its monopoly of power to force people to do things – to go to prison, to surrender their children, to give up their assets. This section focuses on the most important rules used to decide legal trials – the *standards of proof*.

1.1 The Criminal Standard

A standard of proof is a rule used to determine when the evidence is strong enough for a positive judgement (e.g. finding someone guilty of a crime) to be appropriate. There are different standards of proof. One of the most important distinctions is between the standard used in the criminal law and in the civil law. Criminal proof is the primary focus of this Element, but we will discuss civil proof as we go along. All of this is easiest to appreciate at the level of concrete detail, so let's jump in.

Criminal law ranges over conduct that has been criminalised – such as murder, theft, assault, sexual offences, fraud, and so on. Criminal conduct is usually prosecuted *by* the state (rather than the victim) *against* an individual. Criminal law is distinctive because those judged guilty are open to receive the most serious sanctions available to the legal system. These sanctions are *punitive* and can involve the imposition of serious harms on the offender – such as imprisonment or (in some jurisdictions) corporal and capital punishment.

In a criminal trial, the standard of proof used in many jurisdictions is:

Criminal standard of proof = prove guilt *beyond reasonable doubt*[1]

This standard instructs the court to convict the accused of a crime *only if* the evidence supports the guilt of the accused beyond a reasonable doubt. If this standard is not met, the accused must be judged not guilty ('acquitted'). So, the criminal standard specifically governs 'guilty' verdicts, with 'not guilty' verdicts being returned whenever the standard of proof for guilt is not met.

Beyond reasonable doubt ('BRD' for short) is obviously a demanding standard. There are many things supported by good evidence that are nevertheless reasonable to doubt. For example, there may be good reasons to trust your sometimes unreliable friend when they promise to meet you at 7 p.m. for a beer. But it might also be reasonable to harbour doubts. The BRD standard tells us to convict only if there are no reasonable doubts. This means that even if there is some evidence that the accused is guilty – even if you think there is a 'good chance' they are guilty – the court should release the accused so long as there is reasonable doubt. There are much less demanding standards of proof we might use. Indeed, there are other legal standards of proof used outside the trial context. For example, the standard used in different jurisdictions within the United Kingdom to decide whether the police can stop and search somebody is 'reasonable grounds for suspicion'. Clearly, it would be a rather different criminal justice system if criminal courts imprisoned anybody against whom there was reasonable suspicion! Beyond reasonable doubt is a demanding standard of proof.

1.1.1 The *Actus Reus* and *Mens Rea*

Since it will be important later, it is worth saying more about *what* must be proven against the BRD standard to establish criminal guilt. In the 'common law' legal systems we focus on, there are two components jointly required to prove someone has committed a crime.[2] These two components are known by their Latin names – the *actus reus* and the *mens rea.* While the Latin may be unhelpful, the basic idea is pretty straightforward.

First, the *actus reus* is the 'active' part of the crime – the action or conduct that the law prohibits. Here are some rough examples. For theft, the prohibited action is taking the possessions of another without authority, for rape the prohibited action is non-consensual sexual intercourse, for murder the

[1] I focus on Anglo-American systems, which have served as the model for many international institutions (like the International Criminal Court). Other jurisdictions use different phrasings for the criminal standard, but generally have a similarly demanding approach.

[2] Common law legal systems are characteristically defined by reliance on what judges have said in previous cases – 'precedent' – to interpret and create law.

prohibited action is causing the death of another person, and so on. Perhaps obviously, criminal conviction requires proving that the accused performed the *actus reus* – the action (or sometimes omission) that constitutes a crime.[3]

Second, the *mens rea* is the 'mental' component of the crime. For example, 'intent' is a common *mens rea* found in the definition of many crimes. For a crime with a *mens rea* of intent, the *actus reus* has to be performed intentionally. This demonstrates where the *actus reus* and the *mens rea* can diverge; for example, someone can cause a death unintentionally. Something is only a crime when there is the right kind of unity between *actus reus* and *mens rea*, between the action that is performed and the mental state underpinning it. There are other mental states beyond intent that can be the *mens rea* for various crimes and we will come back to these later, but this simple account should be enough to go on for now.

To *prove* a crime, the prosecution has the burden of proof to establish *both* the *actus reus* and the *mens rea*, against the standard of proof. So, take the example of theft. To prove the crime of theft you must show that it is beyond reasonable doubt that the accused took the property *and* that it is beyond reasonable doubt that they intended to do so.

1.2 The Civil Standard

Let us now turn to the second main standard of proof, the one used in the civil law.

The civil law regulates the wide variety of non-criminal disputes adjudicated by the legal system. This includes contractual disputes, employment law, corporate law, family law (e.g. disputes about divorce), disputes about 'negligence' (often called torts), and constitutional law. Civil cases can be pursued by almost any person or legal entity against almost any other person or legal entity. In most civil cases, the standard of proof is the following:

Civil standard of proof = prove your case on the *balance of probabilities*

A common way to think about the balance of probabilities is just that it means 'more likely than not'. So, take an example from employment law. An employer should be held liable for breaching their obligations (e.g. failing to provide safety equipment) just so long as the court finds it more likely than not that they failed to provide safety equipment and had an obligation to provide it.

[3] A criminal omission might occur where a duty to prevent something is imposed by law – for example, as might apply to public office holders. 'Attempts' can also be a criminal *actus reus*, such as attempting to kill somebody. There are tricky issues in determining when someone has committed a criminal attempt rather than just having a vague plan or intention.

The sanctions that result from losing a civil case are extremely varied. The most common is being compelled to pay compensation. Financial penalties can be vast and serious, leading to bankruptcy or impoverishment. But civil cases cannot generally lead to someone being imprisoned or subjected to other types of punitive treatment. A general rule of thumb in civil law is that compensation aims to provide 'restitution' rather than punishment; it aims to put the party that was harmed in as good a place as they would have been had you not harmed them. For example, the court might try to estimate how much your interests were set back by having your contract breached and ask the other side to make up for it in monetary terms.

Clearly, the criminal standard is harder to satisfy than the civil standard – you can reckon something is 'more likely than not' while still having reasonable doubts. I might think it is more likely than not that my sometimes unreliable friend will turn up for our beer, but due to their track record I may have reasonable doubts. It is entirely consistent for evidence to be strong enough to satisfy the civil standard but not the criminal standard. Indeed, there are instances where this happens. Sometimes people are found not guilty of sexual assault in a criminal court (and hence not subject to punitive treatment like imprisonment) but found liable for sexual assault in a civil court (and hence asked to pay monetary compensation).

1.3 Justifying the Standards

How do we come up with the different standards of proof? In truth, the standards of proof have been heavily influenced by historical circumstance and have evolved piecemeal over time. Especially in common law countries where judicial opinions in individual cases influence the way the law evolves – the system of 'precedent' – the history of legal rules is often convoluted rather than the product of a single design. (Of course, convoluted and complex evolutionary processes do not necessarily make for worse products.) Legal history has, for me, a compulsive nerdish attraction because it shows how fragile and often accidental the way that the law works is. As we will see throughout the Element, it is also a source of stories, where idiosyncratic characters find themselves in the courts and change the way that entire states have operated.

The history of the 'beyond reasonable doubt' standard is much discussed, and the introduction of this terminology happened gradually rather than all at once. The language of 'reasonable doubt' evolved from philosophical discussions throughout the 1600s–1800s that worried about the fact that it might be impossible to prove almost anything with absolute (or 'metaphysical') certainty. (Remember Descartes!) Instead, it was thought that proof of everyday matters,

where there is always a chance of error, should be linked to the conscience of the person looking at the evidence; this is sometimes called proving something with 'moral' certainty.[4] From this religiously inflected language, the secular idea of proof beyond reasonable doubt emerged. The connection between proof and the individual conscience will reappear throughout the Element.

As philosophers we are not *primarily* interested in how legal standards came to be the way they are. Rather, we are interested in how they ought to be. We can break this up into two related questions:

(1) Can we reconstruct a justification for the current standards of proof?
(2) In light of how we justify the standards, are they set in the best way?

A natural way to try and justify the different standards of proof is to think about how bad different types of mistakes would be. This is because what standards of proof do, in effect, is strike a balance between different types of error. To see this, consider criminal law.

When criminal courts make decisions – either finding someone guilty or not guilty – they can get it wrong in two different ways. One mistake is *convicting an innocent* person and punishing them for a crime they did not commit. Another type of mistake is *acquitting a guilty* person, allowing a criminal to walk free unpunished. Both of these mistakes are bad, regardless of what type of moral theory you endorse. Convicting the innocent and acquitting the guilty (generally) has bad consequences; the former inflicts misery, while the latter can mean that a dangerous person is released back into the community. It also seems to be unfair even aside from the consequences; the innocent don't *deserve* to be punished, while the guilty might *deserve* punishment. (Of course, this assumes a rather traditional view about the value of punishment. Some might wonder whether punishing the guilty really has good consequences or whether it really is true that those who commit crimes *deserve* to be harmed. While I have some sympathy with this outlook, scepticism about punishment will mostly wait for another day.)

1.4 Two Types of Mistake

Setting the standard of proof involves a balancing act. The harder we make convicting someone of a crime, the less often we will convict people for crimes they didn't commit. Demanding standards provide protection to the accused. However, by making the standard harder to satisfy, we thereby also make it more likely we will acquit people of crimes they *are* guilty of. And, of course, the converse is equally true. The lower the standard, the easier it is to convict; we end up blaming

[4] Shapiro 1986.

more people who are genuinely guilty, but also, we make it more likely that we will convict the innocent. How should we perform this balancing act?

One approach might be to suppose that different decisions have some type of expected value or benefit. Then, once we have thought about how large or small these values are, we try to set the standard in a way that would maximise the expected overall value of all the decisions that courts make.

That's quite abstract, so here's an analogy. Suppose you are a fisherman deciding what net to use. You are only after a certain type of fish, red snappers. They are the only fish you can sell at the market; catching other types of fish drains your resources and harms the fish unnecessarily. If you use a very fine net, you'll let fewer red snappers escape, even snappers that are small and difficult to ensnare. But you'll also catch other types of fish too, ones you don't want. If you use a coarser net, you avoid mistakenly landing the unwanted fish you would catch with the fine net, but you also let more precious red snappers get away. What type of net should you use? Well, it depends on the relative value of catching the snappers compared to the expense of catching the unwanted fish. If there's a big difference in value, then we might be justified in using a very coarse or a very fine net. In the criminal law, in effect, we currently use a very coarse net. We leave aside many finer nets – namely, weaker standards of proof – that would catch more guilty people. But why?

1.5 Blackstone's Asymmetry

The following idea by the English jurist, William Blackstone, is often used by way of justification:

Blackstone's asymmetry: It is much worse to mistakenly convict an innocent person than to mistakenly acquit a guilty person.

Blackstone himself suggested it would be *ten times* worse to convict an innocent than acquit the guilty, famously saying: 'All presumptive evidence of felony should be admitted cautiously, *for the law holds that it is better that ten guilty persons escape than that one innocent suffer.*'[5] It's an interesting historical question why Blackstone's 10:1 ratio became the canonical version of the asymmetry known to all law students. In fact, throughout history we find a dizzying number of attempts to formulate a ratio. Even the Book of Genesis contains a passage in which Moses asks how many innocents would have to be present in Sodom in order to

[5] Blackstone 1827, book four, chapter 27 (emphasis added).

prevent it from being destroyed by God.[6] But something about the 10:1 ratio has a ring of plausibility. There are ways to build formal models designed to maximise the expected utility of our decisions – a branch of philosophy called 'decision theory' – which suggest that Blackstone's 10:1 ratio recommends a 90 per cent level of confidence as the best level at which to set the threshold for conviction.[7] Intriguingly, 90 per cent confidence in guilt is roughly where some people settle when attempting to quantify the BRD standard of proof. (We'll have more on probabilistic approaches to proof later on.) This 90 per cent confidence level also matches up, roughly, with what some empirical surveys have said about how BRD is interpreted by judges.[8]

However, there is a big question facing Blackstone's seductive asymmetry. How can we justify the claim that it is *much worse* to convict an innocent than mistakenly release the guilty?

One way to think about criminal justice is to focus on what people 'deserve' irrespective of the consequences (a view sometimes called '*retributivism*'). For instance, punishing an elderly criminal whose victims have long since died might be expensive and yield little obvious future-oriented benefit. Yet, some might think that seeking to convict such a person is the just thing to do regardless of whether it leads to any particular beneficial consequences. There are different ways to elaborate on this idea of criminal justice aiming to give people what they deserve. But, it isn't immediately obvious that focusing on what people deserve justifies a very demanding standard of proof. While it is true that convicting the innocent fails to give people what they deserve, so does mistakenly acquitting the guilty. Consider the following. One worry, to which we will return repeatedly as a matter of policy interest, is the low conviction rate for sexual crimes. There is a striking drop-off rate in the number of sexual offences reported to the police against the number that are prosecuted in the courtroom. For example, in England and Wales, a recent report claims that less than 2 per cent of complaints lead to a conviction.[9] One possibility is that the high standard of proof is partly responsible for the low conviction rates for certain types of crimes. Proof beyond a reasonable doubt can be difficult in the context of sexual criminality due to the often private nature of such crimes. If

[6] See Volokh 1997 for an entertaining discussion.
[7] On such approaches, see Kaplan 1968; Lillquist 2002. [8] For example, see Solan 1999.
[9] HM Government 2021, 7. See Thomas 2023 for empirical work on the conviction rate for sexual offences at trial: her findings suggest that criminal trials themselves may not be site of the problem, with conviction rates above 65 per cent in recent years. Of course, the lesson of the low complaint-to-conviction ratio is that many allegations never make it to trial. The standards used in trials influence both police and prosecutorial decisions. For instance, prosecutors will decline to pursue a charge precisely because they believe it is unlikely to meet the standard of criminal proof used at trial.